Appendix E: GRADE Profiles

E.1 Information needs of people with Parkinson's disease and their families and carers

E.1.1 Impulse control behaviours

Quality of life impact of having ICD

Quality assess	ment					Number of patients		Effect			
Number of studies	Design	Risk of bias	Inconsiste ncy	Indirectnes s	Imprecisio n	ICD	No ICD	Mean difference / Odds ratio: (95% Cl)	Quality		
Effect of ICD on quality of life (PDQ39)											
Phu (2014)CohortNot seriousN/A1Not serious2Not serious31585MD = 18 (2.24 to 33.76)H											
Patient experie	ence: major de	pressive diso	rder in ICD								
Phu (2014)	Cohort	Not serious	N/A ¹	Not serious ²	Serious ⁴	15	85	OR = 3.07 (0.80 to 11.69)	MODERATE		
¹ N/A: not applica ² No serious indire ³ CI do not cross ⁴ Serious imprecis	 ¹ N/A: not applicable as only one study contributed to this analysis ² No serious indirectness: population matches review protocol ³ CI do not cross MID of 1.6 points (Peto et al., 2001) ⁴ Serious imprecision: Non-significant results 										

Reluctance to start medication for Parkinson's disease

Quality asses	sment										
Number of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Number of patients	Number of physicians	Quality			
A mutual mise	understanding by	y patients and ph	nysicians								
Mestre 2014	Cross- sectional	Serious ¹	N/A ²	Not serious ³	Not serious	62/201	268	MODERATE			
¹ Serious risk of methodology ² N/A: not applic	¹ Serious risk of bias: Methodology not clear, not clear whether all survey/questionnaire materials were standardised or validated as assessed by the reviewer (no well-validated methodology quality checklist available for cross-sectional studies) ² N/A: not applicable as only one study contributed to this analysis										
Mestre 2014 Cross- sectional Serious ¹ N/A ² Not serious ³ Not serious 62/201 268 MO ¹ Serious risk of bias: Methodology not clear, not clear, not clear whether all survey/questionnaire materials were standardised or validated as assessed by the reviewer (no well-val methodology quality checklist available for cross-sectional studies) Not serious ³ Not serious 62/201 268 MO ² N/A: not applicable as only one study contributed to this analysis Not serious ³ Not serious ³ Not serious 62/201 268 MO											

³ No serious indirectness: population matches review protocol

E.1.2 Women of childbearing age

Birth complications in women with PD

Quality assessm	ent									
Studies	Design	Risk of bias	Inconsistency	Indirectness	imprecision	Successful pregnancies	Spontaneous miscarriages in the first 4 months of pregnancy	Quality		
Number of spont	aneous miscar	riages in the firs	t 4 months of pre	egnancy						
Golbe 1987	Qualitative	Very serious ^a	N/A ^b	Not serious ^c	Serious ^d	N= 17	N= 3/17 (15%)	VERY LOW		
Number of total e	elective abortio	ns								
Golbe 1987	Qualitative	Very serious ^a	N/A ^b	Not serious ^c	Serious ^d	N= 17	N= 4/17 (24%)	VERY LOW		
Mean PD disease	e duration									
Golbe 1987	Qualitative	Very serious ^a	N/A ^b	Not serious ^c	Serious ^d	4.2 (4.5) years	3 (2.6) years	VERY LOW		
 ^a Very serious risk of bias as assessed by CASP qualitative quality checklist ^b N/A: not applicable as only one study contributed to this analysis ^c No serious indirectness: population matches review protocol ^d Serious imprecision: Number of participants small 										

Pregnancy complications and related drug therapy in women with PD

Quality assessm								
Studies	Design	Risk of bias	Inconsistency	Indirectness	imprecision	Treatment	No treatment	Quality
Rate of complica								
Golbe 1987	Case series	Very serious ^a	N/A ^b	Not serious ^c	Serious ^d	4/4 (100%) (2 miscarriage, 1 preeclampsia, 1 1 st tri bleeding)	4/16 (25%) (vaginal bleeding or severe nausea)	VERY LOW
Rate of complica	tions associa	ted with leved	ona/carhidona					

ate of complications associated with levodopa/carbidopa

Quality assessm								
Studies	Design	Risk of bias	Inconsistency	Indirectness	imprecision	Treatment	No treatment	Quality
Golbe 1987	Case series	Very serious ^a	N/A ^b	Not serious ^c	Serious ^d	4/6 (66%) (worsening of PD symptoms)	NA	VERY LOW
 ^a Very serious risk of ^b N/A: not applicable ^c No serious indirect ^d Serious imprecision 	of bias as assess e as only one stu tness: population on: Number of pa	ed by CASP qual udy contributed to n matches review articipants small	itative quality check this analysis protocol	list				

Neurological complications of pregnancy in women with PD

Quality assessment									
Example Studies	Design	Risk of bias	Inconsiste ncy	Indirectnes s	imprecision	Total number of pregnancies	Events	Quality	
Exacerbation of PD symp									
Golbe 1987	Case series	Very serious ^a	N/A ^b	Not serious ^c	Serious ^d	17	11/17 (64.7%)	VERY LOW	
Improvement of PD symp	toms post-del	livery (in popu	lation who ex	perienced wor	sening during	pregnancy)			
Golbe 1987	Case series	Very seriousª	N/A ^b	Not serious ^c	Serious ^d	11	1/11 (9.09%)	VERY LOW	
Development of serious p	ost-partum de	pression requ	iring medicat	ion					
Golbe 1987	Case series	Very seriousª	N/A ^b	Not serious ^c	Serious ^d	4	0/4 (0%)	VERY LOW	
 ^a Very serious risk of bias as assessed by CASP qualitative quality checklist ^b N/A: not applicable as only one study contributed to this analysis ^c No serious indirectness: population matches review protocol ^d Serious imprecision: Number of participants small 									

Post-partum depression/anxiety

Quality assessment Number of patients								
Example Studies	Design	Risk of bias	Inconsiste ncy	Indirectnes s	imprecision	Total number of pregnancies	Events	Quality
Development of serious p								
Golbe 1987	Case series	Very serious ^a	N/A ^b	Not serious ^c	Serious ^d	4	0/4 (0%)	VERY LOW
 ^a Very serious risk of bias as assessed by CASP qualitative quality checklist ^b N/A: not applicable as only one study contributed to this analysis ^c No serious indirectness: population matches review protocol ^d Serious imprecision: Number of participants small 								

4

E.2 Pharmacological management of motor symptoms

E.2.1 First-line treatment of motor symptoms

E.2.1.1 Treatment-naïve population

UPDRS Total – MAOB (Rasagiline, Selegiline) vs. placebo

Change in UPDRS Total from baseline to 36 weeks/12 months – MAOB vs. placebo

Quality assessment	Quality assessment							Effect			
		Risk of							Qualit		
Number of studies	Design	bias	Inconsistency	Indirectness	Imprecision	Intervention	Control	Mean Difference (95% CI)	У		
Change in UPDRS total score											
2 studies:	RCT	Serious ¹	Not serious	Serious ⁵	Not serious ⁶	613	612	-3.07 (-3.78, -2.37)	LOW		
Olanow et al., 2009;											
Palhågen et al., 1998											
 ¹ Downgraded 1 level: Serious risk of bias as assessed by NICE RCT quality checklist ² N/A: Not applicable, only 1 study contributed to this analysis ³ No serious indirectness: population was as specified in review protocol ⁴ Downgraded 1 level: Non-significant results 											
⁵ Downgraded 1 level: Some	patients from	m the placebo grou	up had early transfer f	rom phase 1 to pha	se 2, where active	treatment was give	n				
6 CI do not cross MID of 7.3	ooints (Schr	ag et al., 2006)									

Beck Depression Inventory - Pramipexole vs. placebo

BDI from baseline to 9 months – Pramipexole vs. placebo

Quality assessment				Number of patients		Effect	Quality				
Number of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Intervention	Control	Mean Difference (95% CI)			
Change in BDI score											
1 study: Schapira et al., 2013	RCT	Serious ¹	N/A ²	Serious ⁵	Not serious	211	200	-1.4 (-2.23, -0.57)	LOW		
 ¹ Downgraded 1 level: Serious risk of bias as assessed by NICE RCT quality checklist ² N/A: Not applicable, only 1 study contributed to this analysis ³ No serious indirectness: population was as specified in review protocol ⁴ Downgraded 1 level: Non-significant results 											

Quality assessment			Number of pat	ients	Effect	Quality			
Number of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Intervention	Control	Mean Difference (95% CI)	

⁵ Downgraded 1 level: Some patients from the placebo group had early transfer from phase 1 to phase 2, where active treatment was given

⁶ CI do not cross MID of 7.3 points (Schrag et al., 2006)

Adverse events - Ropinirole vs. Pramipexole (dopamine agonists)

Any AE leading to trial discontinuation - Ropinirole vs. pramipexole

Quality assessment			Number of patients		Effect	Quality			
Number of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Intervention	Control	RR (95% CI)	
Adverse event									
1 study: Thomas et al., 2006	RCT	Serious ¹	N/A ²	Not serious ³	Serious ⁴	30	30	1.67 (0.44, 6.36)	LOW
 Downgraded 1 level: Seriou: N/A: Not applicable, only 1 s No serious indirectness: pop Downgraded 1 level: Non-si Downgraded 1 level: Some CI do not cross MID of 7.3 p 	s risk of bias study contribu- pulation was gnificant resu- patients from points (Schra	as assessed by NIC uted to this analysis as specified in revier ults n the placebo group I g et al., 2006)	E RCT quality checklist w protocol nad early transfer from p	hase 1 to phase 2, w	here active treatmen	t was given			

Adverse events - Rasagiline vs. placebo

Adverse event rate (any AE) - Rasagiline vs. placebo

Quality assessment				Number of patients		Effect	Quality				
Number of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Intervention	Control	IRR (95% CI)			
Adverse event rate											
1 study:	RCT	Serious ¹	N/A ²	Serious ⁵	Not serious	576	588	0.80 (0.65, 0.99)	LOW		
Olanow et al., 2009											
 ¹ Downgraded 1 level: Serious ² N/A: Not applicable, only 1 s ³ No serious indirectness: pop ⁴ Downgraded 1 level: Non-sig 	 ¹ Downgraded 1 level: Serious risk of bias as assessed by NICE RCT quality checklist ² N/A: Not applicable, only 1 study contributed to this analysis ³ No serious indirectness: population was as specified in review protocol ⁴ Downgraded 1 level: Non-significant results 										
⁵ Downgraded 1 level: Some patients from the placebo group had early transfer from phase 1 to phase 2, where active treatment was given											
⁶ CI do not cross MID of 7.3 p	oints (Schra	ig et al., 2006)									

Adverse event rate (AE related to dopaminergic therapy) – Rasagiline vs. placebo

Quality assessment	Quality assessment							Effect	Quality	
Number of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Intervention	Control	IRR (95% CI)		
Adverse event rate										
1 study:	RCT	Serious ¹	N/A ²	Serious ⁵	Serious ⁴	576	588	0.72 (0.49, 1.07)	VERY LOW	
Olanow et al., 2009	iow et al., 2009									
¹ Downgraded 1 level: Serior ² N/A: Not applicable, only 1 ³ No serious indirectness: pc ⁴ Downgraded 1 level: Non-s	us risk of bia study contri pulation was significant re	is as assessed by N buted to this analysi s as specified in rev sults	ICE RCT quality check s iew protocol	list						
⁵ Downgraded 1 level: Some	³ Downgraded 1 level: Some patients from the placebo group had early transfer from phase 1 to phase 2, where active treatment was given									
⁶ CI do not cross MID of 7.3	points (Schr	ag et al., 2006)								

Adverse events - Levodopa/carbidopa (150/37.5 mg/day and 300/75 mg/day) vs. placebo

Adverse event rate (any AE) - Levodopa/carbidopa (150/37.5 mg/d and 300/75 mg/day) vs. placebo

Quality assessment	Number of patients		Effect	Quality					
Number of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Intervention	Control	IRR (95% CI)	
Adverse event rate									
1 study: Fahn et al., 2005	RCT	Serious ¹	N/A ²	Not serious ³	Serious ⁴	180	90	1.00 (0.84, 1.20)	LOW
 ¹ Downgraded 1 level: Seriou ² N/A: Not applicable, only 1 state ³ No serious indirectness: pop ⁴ Downgraded 1 level: Non-sis ⁵ Downgraded 1 level: Some 	is risk of bias study contrib pulation was ignificant res patients fron	as assessed by NIC uted to this analysis as specified in revie ults n the placebo group	CE RCT quality checklist w protocol had early transfer from p	ohase 1 to phase 2, w	here active treatmen	it was given			

⁶ CI do not cross MID of 7.3 points (Schrag et al., 2006)

Adverse event rate (AE related to dopaminergic therapy) - Levodopa/carbidopa (150/37.5 mg/d and 300/75 mg/day) vs. placebo

					<u> </u>				
Quality assessment	Quality assessment							Effect	Quality
Number of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Intervention	Control	IRR (95% CI)	
Adverse event rate									
1 study: Fahn et al., 2005	RCT	Serious ¹	N/A ²	Not serious ³	Serious ⁴	180	90	0.85 (0.60, 1.21)	LOW
¹ Downgraded 1 level: Seriou ² N/A: Not applicable only 1 s	s risk of bias	as assessed by NIC	CE RCT quality checklist						

Quality assessment	Quality assessment							Effect	Quality
Number of studies	nber of studies Design Risk of bias Inconsistency Indirectness Imprecision						Control	IRR (95% CI)	
³ No serious indirectness: pop	ulation was	as specified in review	w protocol						

⁴ Downgraded 1 level: Non-significant results

⁵ Downgraded 1 level: Some patients from the placebo group had early transfer from phase 1 to phase 2, where active treatment was given

⁶ CI do not cross MID of 7.3 points (Schrag et al., 2006)

Serious adverse event rate - Levodopa/carbidopa (150/37.5 mg/d and 300/75 mg/day) vs. placebo

Quality assessment	Quality assessment							Effect	Quality		
Number of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Intervention	Control	IRR (95% CI)			
Adverse event rate											
1 study:	RCT	Serious ¹	N/A ²	Not serious ³	Serious ⁴	180	90	1.50 (0.41, 5.54)	LOW		
Fahn et al., 2005	., 2005										
 Downgraded 1 level: Seriou: N/A: Not applicable, only 1 s No serious indirectness: pop Downgraded 1 level: Non-si 	 ¹ Downgraded 1 level: Serious risk of bias as assessed by NICE RCT quality checklist ² N/A: Not applicable, only 1 study contributed to this analysis ³ No serious indirectness: population was as specified in review protocol ⁴ Downgraded 1 level: Non-significant results 										
⁵ Downgraded 1 level: Some patients from the placebo group had early transfer from phase 1 to phase 2, where active treatment was given											
⁶ CI do not cross MID of 7.3 p	CI do not cross MID of 7.3 points (Schrag et al., 2006)										

Adverse events - Levodopa/cabidopa (600/150 mg/day) vs. placebo

Adverse event rate (any AE) - Levodopa/carbidopa (600/150 mg/day) vs. placebo

Quality assessment	Quality assessment							Effect	Quality
Number of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Intervention	Control	IRR (95% CI)	
Adverse event rate									
1 study: Fahn et al., 2005	RCT	Serious ¹	N/A ²	Not serious ³	Serious ⁴	91	90	1.18 (0.97, 1.43)	LOW
 ¹ Downgraded 1 level: Serious ² N/A: Not applicable, only 1 s ³ No serious indirectness: pop ⁴ Downgraded 1 level: Non-sig ⁵ Downgraded 1 level: Some ⁶ Downgraded 1 level: Some 	s risk of bias study contrib pulation was gnificant resu patients from	as assessed by NIC uted to this analysis as specified in revie ults the placebo group	E RCT quality checklist w protocol had early transfer from p	phase 1 to phase 2, w	here active treatmen	t was given			
⁶ CI do not cross MID of 7.3 p	oints (Schra	ig et al., 2006)							

Adverse event rate (AE related to dopaminergic therapy) – Levodopa/carbidopa (600/150 mg/day) vs. placebo

Quality assessment	Quality assessment							Effect	Quality
Number of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Intervention	Control	IRR (95% CI)	
Adverse event rate									
1 study: Fahn et al., 2005	RCT	Serious ¹	N/A ²	Not serious ³	Serious ⁴	91	90	1.23 (0.84, 1.78)	LOW
 ¹ Downgraded 1 level: Seriou ² N/A: Not applicable, only 1 s ³ No serious indirectness: pop ⁴ Downgraded 1 level: Non-si ⁵ Downgraded 1 level: Some ⁶ Old each series MID of 7 2 m 	s risk of bias study contrib pulation was gnificant resp patients from	as assessed by NIC uted to this analysis as specified in revie- ults the placebo group I	E RCT quality checklist w protocol nad early transfer from p	hase 1 to phase 2, w	here active treatmen	t was given			

⁶ CI do not cross MID of 7.3 points (Schrag et al., 2006)

Serious adverse event rate - Levodopa/carbidopa (600/150 mg/day) vs. placebo

Quality assessment	Number of patients		Effect	Quality					
Number of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Intervention	Control	IRR (95% CI)	
Adverse event rate									
1 study: Fahn et al., 2005	RCT	Serious ¹	N/A ²	Not serious ³	Serious ⁴	91	90	0.66 (0.11, 3.95)	LOW
¹ Downgraded 1 level: Serious ² N/A: Not applicable, only 1 s ³ No serious indirectness: pop ⁴ Downgraded 1 level: Non-sig ⁵ Downgraded 1 level: Some p	s risk of bias study contribu- pulation was gnificant response from patients from	as assessed by NIC uted to this analysis as specified in revier ults n the placebo group I in the placebo group I	E RCT quality checklist w protocol nad early transfer from p	phase 1 to phase 2, w	here active treatmen	t was given			

Adverse events - Pramipexole vs. placebo

Any adverse event - Pramipexole vs. placebo

Quality assessment	·			Number of patients		Effect	Quality		
Number of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Intervention	Control	RR (95% CI)	
Adverse event									
1 study:	RCT	Not serious	N/A ²	Serious ⁵	Serious ⁴	261	274	1.04 (0.94, 1.15)	LOW
Schapira et al., 2013									
¹ Downgraded 1 level: Serious ² N/A: Not applicable, only 1 st	risk of bias a udy contribu	as assessed by NICI ted to this analysis	E RCT quality checklist						

Quality assessment	Quality assessment						ents	Effect	Quality
Number of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Intervention	Control	RR (95% CI)	
³ No serious indirectness: population was as specified in review protocol									

⁴ Downgraded 1 level: Non-significant results

⁵ Downgraded 1 level: Some patients from the placebo group had early transfer from phase 1 to phase 2, where active treatment was given

⁶ CI do not cross MID of 7.3 points (Schrag et al., 2006)

Any serious adverse event - Pramipexole vs. placebo

Quality assessment	Number of patients		Effect	Quality						
Number of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Intervention	Control	RR (95% CI)		
Adverse event										
1 study:	RCT	Not serious	N/A ²	Serious ⁵	Serious ⁴	261	274	0.99 (0.52, 1.88)	LOW	
Schapira et al., 2013	ipira et al., 2013									
 Downgraded 1 level: Serious N/A: Not applicable, only 1 st No serious indirectness: popu Downgraded 1 level: Non-sig 	 ¹ Downgraded 1 level: Serious risk of bias as assessed by NICE RCT quality checklist ² N/A: Not applicable, only 1 study contributed to this analysis ³ No serious indirectness: population was as specified in review protocol ⁴ Downgraded 1 level: Non-significant results 									
⁵ Downgraded 1 level: Some patients from the placebo group had early transfer from phase 1 to phase 2, where active treatment was given										
⁶ CI do not cross MID of 7.3 pc	CI do not cross MID of 7.3 points (Schrag et al., 2006)									

Any serious adverse event leading to discontinuation - Pramipexole vs. placebo

Quality assessment	Quality assessment							Effect	Quality
Number of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Intervention	Control	RR (95% CI)	
Adverse event									
1 study:	RCT	Not serious	N/A ²	Serious ⁵	Serious ⁴	261	274	1.01 (0.60, 1.70)	LOW
Schapira et al., 2013									
¹ Downgraded 1 level: Serious	risk of bias a	as assessed by NIC	E RCT quality checklist						
² N/A: Not applicable, only 1 st	udy contribu	ted to this analysis							
³ No serious indirectness: population was as specified in review protocol									
⁴ Downgraded 1 level: Non-sig	nificant resul	Its							

⁵ Downgraded 1 level: Some patients from the placebo group had early transfer from phase 1 to phase 2, where active treatment was given

⁶ CI do not cross MID of 7.3 points (Schrag et al., 2006)

Network meta-analyses

UPDRS Total

Quality assessment									
Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Quality				
Change in UPDRS Total score									
5 MAOB: Mally et al., 1995; Palhågen et al., 1998; Olanow et al., 2009. DA: Schapira et al., 2013. Levodopa: Fahn et al., 2005.	5 Serious ¹ Not serious Not serious ³ Not serious MAOB: Mally et al., 1995; Palhågen et al., 1998; Olanow et al., 2009. DA: Schapira et al., 2013. Levodopa: Fahn et al., 2005.								
¹ Downgrade 1 level: Limitations in the design or execution of the study ² No heterogeneity (i ² =0%) ³ Considered not serious as population, interventions, comparator and outcomes are as defined in protocol									

UPDRS II (ADL)

Quality assessment									
Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Quality				
Change in UPDRS ADL score									
4 MAOB: Mally et al., 1995; Palhågen et al., 1998. DA: Schapira et al., 2013. Levodopa: Fahn et al., 2005.	Serious ¹	Serious ²	Not serious ³	Not serious	LOW				
¹ Downgrade 1 level: Limitations in the design or execution of the study ² Considerable between study heterogeneity (i ² >40%) ³ Considered not serious as population, interventions, comparator and outcomes are as defined in protocol									

UPDRS III (Motor)

Quality assessment							
Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Quality		
Change in UPDRS Motor score							
4	Serious ¹	Serious ²	Not serious ³	Not serious	LOW		

Quality assessment								
Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Quality			
MAOB: Mally et al., 1995; Palhågen et al., 1998. DA: Schapira et al., 2013. Levodopa: Fahn et al., 2005.								
¹ Downgrade 1 level: Limitations in the design or execution of the study ² Considerable between study heterogeneity (i ² >40%) ³ Considered not serious as population, interventions, comparator and outcomes are as defined in protocol								

E.2.1.2 Full population

Low-dose levodopa versus placebo

Outcome	No. of studies	Risk of bias	Inconsistency	Indirectness	Imprecision	Estimate (CI)	Overall quality
UPDRS (ADL)	1 (Fahn)	Serious ¹	N/A	Not serious	Not serious ⁵	MD -1.60 (-2.64, -0.56)	Moderate
UPDRS (motor)	1 (Fahn)	Serious ¹	N/A	Not serious	Serious ³	MD -2.90 (-4.94, -0.86)	Low
UPDRS (total)	1 (Fahn)	Serious ¹	N/A	Not serious	Serious ⁴	MD -5.00 (-7.76, -2.24)	Low
Any AE	1 (Fahn)	Serious ¹	N/A	Not serious	Serious ²	IRR 1.01 (0.84, 1.20)	Low
SAE	1 (Fahn)	Serious ¹	N/A	Not serious	Serious ²	IRR 1.50 (0.41, 5.54)	Low
Dopaminergic AE	1 (Fahn)	Serious ¹	N/A	Not serious	Serious ²	IRR 0.85 (0.60, 1.21)	Low

¹Study at high risk of bias; ²Non-significant result; ³CI cross MID: between 3.25 (Horváth et al., 2015) and 5 points (Schrag et al., 2006); ⁴CI cross MID of 7.3 points (Schrag et al., 2006); ⁵CI do not cross MID of 3 points (Schrag et al., 2006)

High-dose levodopa versus placebo

Outcome	No. of studies	Risk of bias	Inconsistency	Indirectness	Imprecision	Estimate (CI)	Overall quality
UPDRS (ADL)	1 (Fahn)	Serious ¹	N/A	Not serious	Serious ³	MD -2.20 (-3.41, -0.99)	Low
UPDRS (motor)	1 (Fahn)	Serious ¹	N/A	Not serious	Serious ⁴	MD -5.40 (-7.85, -2.95)	Low
UPDRS (total)	1 (Fahn)	Serious ¹	N/A	Not serious	Serious ⁵	MD -8.00 (-11.25, -4.75)	Low
Any AE	1 (Fahn)	Serious ¹	N/A	Not serious	Serious ²	IRR 1.18	Low

Outcome	No. of studies	Risk of bias	Inconsistency	Indirectness	Imprecision	Estimate (CI)	Overall quality		
						(0.97, 1.43)			
SAE	1 (Fahn)	Serious ¹	N/A	Not serious	Serious ²	IRR 0.66 (0.11, 3.95)	Low		
Dopaminergic AE	1 (Fahn)	Serious ¹	N/A	Not serious	Serious ²	IRR 1.23 (0.85, 1.78)	Low		
¹ Study at high risk of bias; ² Non-significant result; ³ CI cross MID of 3 points (Schrag et al., 2006); ⁴ CI cross MID: between 3.25 (Horváth et al., 2015) and 5 points (Schrag et al., 2006); ⁵ CI cross MID of 7.3 points (Schrag et al., 2006)									

Extended-release levodopa versus placebo

Outcome	No. of studies	Risk of bias	Inconsistency	Indirectness	Imprecision	Estimate (CI)	Overall quality
UPDRS (ADL)	1 (Pahwa)	Not serious	N/A	Serious ¹	Not serious ⁴	MD -9.23 (-11.61, -6.85)	Moderate
UPDRS (motor)	1 (Pahwa)	Not serious	N/A	Serious ¹	Not serious ⁵	MD -9.23 (-11.61, -6.85)	Moderate
PDQ-39	1 (Pahwa)	Not serious	N/A	Serious ¹	Not serious ⁶	MD -5.31 (-8.90, -1.73)	Moderate
Any AE	1 (Pahwa)	Serious ²	N/A	Serious ¹	Serious ³	RR 0.92 (0.79, 1.06)	Very low
AE discontinuation	1 (Pahwa)	Serious ²	N/A	Serious ¹	Serious ³	RR 2.74 (1.00, 7.52)	Very low

¹Population not treatment-naïve; ²Selection of adverse events to report unclear; ³Non-significant result; ⁴CI do not cross MID of 3 points (Schrag et al., 2006); ⁵CI do not cross MID: between 3.25 (Horváth et al., 2015) and 5 points (Schrag et al., 2006); ⁶CI do not cross MID of 1.6 points (Peto et al., 2001)

Dopamine agonists versus placebo

Short-term follow-up (≤6 months)

Outcome	No. of studies	Risk of bias	Inconsistency	Indirectness	Imprecision	Estimate (CI)	Overall quality
UPDRS (ADL)	6 (Hauser, Hubble, Jankovic, Mizuno, PSG 2003, Zhang)	Not serious ¹	Not serious	Serious ²	Not serious ⁶	MD -1.22 (-1.62, -0.81)	Moderate
UPDRS (motor)	6 (Hauser, Hubble, Jankovic, Mizuno, PSG 2003, Zhang)	Not serious ¹	Not serious	Serious ²	Serious⁵	MD -3.20 (-4.08, -2.31)	Low
UPDRS (total)	2 (Adler, PSG 1997)	Not serious ¹	Not serious	Serious ²	Not serious ⁷	MD -4.85 (-6.65, -3.06)	Moderate
Epworth sleep scale	2 (Hauser, Jankovic)	Not serious	Serious ³	Serious ²	Not serious	MD 1.40 (0.59, 2.22)	Low
PDQ-39	1 (Hauser)	Not serious	N/A	Serious ²	Not serious ⁸	MD -6.81 (-11.42, -2.20)	Moderate
EQ-VAS	1 (Hauser)	Not serious	N/A	Serious ²	Serious ⁴	MD 4.86 (-1.11, 10.84)	Low

¹Individual studies at risk of bias, but overall risk of bias rated low due to consistency of effect between studies at high and low risk of bias; ²Population not treatment-naïve; ³Considerable between study heterogeneity (i² >40%); ⁴Non-significant result; ⁵CI cross MID: between 3.25 (Horváth et al., 2015) and 5 points (Schrag et al., 2006); ⁶CI do not cross MID of 3 points (Schrag et al., 2006); ⁷CI do not cross MID of 7.3 points (Schrag et al., 2006); ⁸CI do not cross MID of 1.6 points (Peto et al., 2001)

Medium term follow-up (6 months – 2.5 years)

Outcome	No. of studies	Risk of bias	Inconsistency	Indirectness	Imprecision	Estimate (CI)	Overall quality
UPDRS (ADL)	2 (Poewe, Schapira)	Not serious	Serious ¹	Serious ²	Not serious ⁶	MD -1.54 (-2.47, -0.62)	Low
UPDRS (motor)	2 (Poewe, Schapira)	Not serious	Serious ¹	Serious ²	Serious ⁴	MD -4.19 (-6.00,2.38)	Very low
UPDRS (total)	1 (Schapira)	Not serious	N/A	Not serious	Not serious ⁷	MD -4.80 (-6.46, -3.14)	High
BDI	1 (Schapira)	Not serious	N/A	Not serious	Not serious	MD -1.40 (-2.23, -0.57)	High
PDQ-39	1 (Poewe)	Not serious	N/A	Serious ²	Serious ⁵	MD -3.63 (-7.01, -0.25)	Low
EQ-VAS	1 (Poewe)	Not serious	N/A	Serious ²	Serious ³	MD 2.94 (-1.46, 7.34)	Low

¹Considerable between study heterogeneity (i² >40%); ²Population not treatment-naïve; ³Non-significant result; ⁴CI cross MID: between 3.25 (Horváth et al., 2015) and 5 points (Schrag et al., 2006); ⁵CI cross MID of 1.6 points (Peto et al., 2001); ⁶CI do not cross MID of 3 points (Schrag et al., 2006); ⁷CI do not cross MID of 7.3 points (Schrag et al., 2006)

Adverse events

Outcome	No. of studies	Risk of bias	Inconsistency	Indirectness	Imprecision	Estimate (CI)	Overall quality
Any AE (Pramipexole)	5 (Hauser, Hubble, Poewe, PSG 1997, Schapira)	Serious ¹	Not serious	Serious ²	Serious ³	RR 1.05 (1.00,1.14)	Very low
Any AE (Rotigotine)	5 (Giladi, Jankovic, PSG 2003, Watts, Zhang)	Serious ¹	Serious ⁴	Serious ²	Not serious	IRR 1.44 (1.09, 1.90)	Very low

Outcome	No. of studies	Risk of bias	Inconsistency	Indirectness	Imprecision	Estimate (CI)	Overall quality
Any AE (Ropinirole)	1 (Adler)	Serious ¹	N/A	Serious ²	Serious ³	RR 1.06 (0.99, 1.13)	Very low
SAE (Pramipexole)	3 (Hauser, Poewe, Schapira)	Serious ¹	Not serious	Serious ²	Serious ³	RR 1.24 (0.74, 2.06)	Very low
SAE (Rotigotine)	2 (Giladi, PSG 2007)	Serious ¹	Not serious	Serious ²	Serious ³	IRR 1.41 (0.68, 2.92)	Very low
SAE (Ropinirole)	1 (Giladi)	Serious ¹	N/A	Serious ²	Serious ³	IRR 1.73 (0.82, 3.63)	Very low
Dopaminergic AE (Pramipexole)	1 (Olanow)	Serious ¹	N/A	Not serious	Serious ³	IRR 1.20 (0.86, 1.67)	Very low
AE discontinuation (Pramipexole)	3 (Hauser, Poewe, Schapira)	Serious ¹	Serious ⁴	Serious ²	Serious ³	RR 0.36 (0.02, 5.97)	Very low
AE discontinuation (Rotigotine)	3 (Giladi, Watts, Zhang)	Serious ¹	Serious ⁴	Serious ²	Not serious	RR 2.07 (1.23, 3.48)	Very low
AE discontinuation (Ropinirole)	2 (Adler, Giladi)	Serious ¹	Not serious	Serious ²	Not serious	RR 2.35 (1.43, 3.86)	Low
¹ Selection of adverse	events to repor	t unclear; ² Populat	tion not treatment-naï	ve; ³ Non-significant	result; ⁴ Considerab	le between study hete	rogeneity (i ² >40%)

Monoamine oxidase inhibitors versus placebo

Short-term follow-up

Outcome	No. of studies	Risk of bias	Inconsistency	Indirectness	Imprecision	Estimate (CI)	Overall quality
UPDRS (ADL)	3 (Mally, Palhågen, PSG 2002)	Not serious ¹	Not serious	Serious ²	Not serious ⁷	MD -1.14 (-1.57, -0.71)	Moderate

Outcome	No. of studies	Risk of bias	Inconsistency	Indirectness	Imprecision	Estimate (CI)	Overall quality
UPDRS (motor)	3 (Mally, Palhågen, PSG 2002)	Not serious ¹	Serious ³	Serious ²	Serious ⁵	MD -4.37 (-7.52, -1.23)	Very low
UPDRS (total)	3 (Hubble, Mally, Palhågen)	Not serious ¹	Serious ³	Serious ²	Serious ⁶	MD -6.38 (-12.33, -0.43)	Very low
BDI	1 (PSG 2002)	Not serious	N/A	Serious ²	Serious ⁴	MD -0.28 (-0.72, 0.16)	Low
PDQUALIF	1 (PSG 2002)	Not serious	N/A	Serious ²	Not serious	MD -2.83 (-3.06, -2.59)	Moderate

¹Individual studies at risk of bias, but overall risk of bias rated low due to consistency of effect between studies at high and low risk of bias; ²Population not treatment-naïve; ³Considerable between study heterogeneity (i² >40%); ⁴Non-significant result; ⁵CI cross MID: between 3.25 (Horváth et al., 2015) and 5 points (Schrag et al., 2006); ⁶CI cross MID of 7.3 points (Schrag et al., 2006); ⁷CI do not cross MID of 3 points (Schrag et al., 2006)

Medium term follow-up (6 months – 2.5 years)

Outcome	No. of studies	Risk of bias	Inconsistency	Indirectness	Imprecision	Estimate (CI)	Overall quality
UPDRS (ADL)	1 (Palhågen)	Serious ¹	N/A	Not serious	Not serious ³	MD -0.30 (-1.50, 0.90)	Moderate
UPDRS (motor)	1 (Palhågen)	Serious ¹	N/A	Not serious	Serious ²	MD -1.90 (-5.26, 1.46)	Low
UPDRS (total)	2 (Olanow, Palhågen)	Serious ¹	Not serious	Not serious	Not serious ⁴	MD -3.07 (-3.78, -2.37)	Moderate

¹Included studies at high risk of bias; ²CI cross MID: between 3.25 (Horváth et al., 2015) and 5 points (Schrag et al., 2006); ³CI do not cross MID of 3 points (Schrag et al., 2006); ⁴CI do not cross MID of 7.3 points (Schrag et al., 2006)

Adverse events

Outcome	No. of studies	Risk of bias	Inconsistency	Indirectness	Imprecision	Estimate (CI)	Overall quality		
Any AE (Rasagiline)	2 (Olanow, Stern)	Serious ¹	Not serious	Serious ²	Not serious	IRR 0.82 (0.68, 1.00)	Low		
SAE (Rasagiline)	1 (PSG 2002)	Serious ¹	N/A	Serious ²	Serious ²	RR 2.08 (0.71, 6.09)	Very low		
Dopaminergic AE (Rasagiline)	1 (Olanow)	Serious ¹	N/A	Not serious	Serious ²	IRR 0.72 (0.49, 1.07)	Low		
¹ Selection of adverse events to report unclear; ² Population not treatment-naïve; ³ Non-significant result									

Levodopa versus dopamine agonists

Outcome	No. of studies	Risk of bias	Inconsistency	Indirectness	Imprecision	Estimate (CI)	Overall quality
UPDRS (ADL)	1 (Holloway)	Not serious	N/A	Serious ¹	Not serious ⁷	MD -1.10 (-1.98, -0.22)	Moderate
UPDRS (motor) - short	1 (Rascol)	Not serious	N/A	Serious ¹	Serious ⁵	MD -2.60 (-4.22, -0.98)	Low
UPDRS (motor)	2 (Holloway, Whone)	Not serious	Not serious	Serious ¹	Serious ⁵	MD -4.69 (-6.29, -3.10)	Low
UPDRS (total)	1 (Holloway)	Not serious	N/A	Serious ¹	Serious ⁶	MD -4.70 (-7.36, -2.04)	Low
Dyskinesia RR	1 (Whone)	No serious	N/A	Serious ¹	Not serious	RR 7.73 (2.39, 25.00)	Moderate
Any AE (Pramipexole)	1 (Holloway)	Serious ²	N/A	Serious ¹	Not serious	IRR 0.55 (0.43, 0.70)	Low
Any AE (Ropinirole)	1 (Rascol)	Serious ²	N/A	Serious ¹	Serious ³	IRR 0.97 (0.84, 1.11)	Very low
SAE (Pramipexole)	1 (Holloway)	Serious ²	N/A	Serious ¹	Serious ³	IRR 0.40 (0.08, 2.08)	Very low

Outcome	No. of studies	Risk of bias	Inconsistency	Indirectness	Imprecision	Estimate (CI)	Overall quality
SAE (Ropinirole)	2 (Rascol, Whone)	Serious ²	Not serious	Serious ¹	Serious ³	RR 1.11 (0.69, 1.80)	Very low
AE discontinuation (Ropinirole)	2 (Rascol, Whone)	Serious ²	Serious ⁴	Serious ¹	Serious ³	RR 0.73 (0.22, 2.39)	Very low

¹Population not treatment-naïve; ²Selection of adverse events to report unclear; ³Non-significant result; ⁴Considerable between study heterogeneity (i² >40%); ⁵CI cross MID: between 3.25 (Horváth et al., 2015) and 5 points (Schrag et al., 2006); ⁶CI cross MID of 7.3 points (Schrag et al., 2006); ⁷CI do not cross MID of 3 points (Schrag et al., 2006); ⁸CI cross MID of 7.3 points (Schrag et al., 2006); ⁹CI do not cross MID of 3 points (Schrag et al., 2006); ⁹CI cross MID of 7.3 points (Schrag et al., 2006); ⁹CI do not cross MID of 3 points (Schrag et al., 2006); ⁹CI cross MID of 7.3 points (Schrag et al., 2006); ⁹CI cross MID of 7.3 points (Schrag et al., 2006); ⁹CI do not cross MID of 3 points (Schrag et al., 2006); ⁹CI cross MID of 7.3 points (Schrag et al., 2006); ⁹CI cross MID of

Long-term follow-up (>2.5 years)

Outcome	No. of studies	Risk of bias	Inconsistency	Indirectness	Imprecision	Estimate (CI)	Overall quality
UPDRS (ADL)	2 (Holloway, Rascol)	Not serious	N/A	Serious ¹	Not serious ⁴	MD -1.32 (-2.28, -0.35)	Moderate
UPDRS (motor)	2 (Holloway, Rascol)	Not serious	N/A	Serious ¹	Serious ²	MD -4.39 (-6.55, -2.23)	Low
UPDRS (total)	2 (Holloway, Rascol)	Not serious	N/A	Serious ¹	Serious ³	MD -5.20 (-8.90, -1.50)	Low
Dyskinesia	2 (Holloway, Rascol)	Not serious	N/A	Serious ¹	Not serious	RR 2.22 (1.74, 2.82)	Moderate

¹Population not treatment-naïve; ²Cl cross MID: between 3.25 (Horváth et al., 2015) and 5 points (Schrag et al., 2006); ³Cl cross MID of 7.3 points (Schrag et al., 2006); ⁴Cl do not cross MID of 3 points (Schrag et al., 2006)

Levodopa versus monoamine oxidase inhibitors

Outcome	No. of studies	Risk of bias	Inconsistency	Indirectness	Imprecision	Estimate (CI)	Overall quality
UPDRS (ADL) - short	1 (Caraceni)	Serious ¹	N/A	Serious ²	Not serious ³	MD -1.10 (-1.62, -0.58)	Low

Outcome	No. of studies	Risk of bias	Inconsistency	Indirectness	Imprecision	Estimate (CI)	Overall quality
UPDRS (motor) - short	1 (Caraceni)	Serious ¹	N/A	Serious ²	Not serious ⁴	MD -1.00 (-2.07, 0.07)	Low

¹Included studies at high risk of bias; ²Population not treatment-naïve; ³CI do not cross MID of 3 points (Schrag et al., 2006); ⁴CI do not cross MID: between 3.25 (Horváth et al., 2015) and 5 points (Schrag et al., 2006)

Long-term follow-up (>2.5 years)

Outcome	No. of studies	Risk of bias	Inconsistency	Indirectness	Imprecision	Estimate (CI)	Overall quality	
Need for add-on therapy	1 (Caraceni)	Not serious	N/A	Serious ¹	Not serious	RR 0.20 (0.13, 0.31)	Moderate	
Motor fluctuations	1 (Caraceni)	Not serious	N/A	Serious ¹	Not serious	RR 1.58 (1.05, 2.37)	Moderate	
Dyskinesia	1 (Caraceni)	Not serious	N/A	Serious ¹	Serious ²	RR 1.30 (0.87, 1.95)	Low	
¹ Population not treatment-naïve: ² Non-significant result								

Dopamine agonists versus monoamine oxidase inhibitors

Outcome	No. of studies	Risk of bias	Inconsistency	Indirectness	Imprecision	Estimate (CI)	Overall quality
Epworth sleep scale	1 (Viallet)	Not serious	N/A	Serious ²	Not serious	MD 1.92 (0.73, 3.11)	Moderate
Any AE (Pramipexole- Rasagiline)	1 (Viallet)	Serious ¹	N/A	Serious ²	Serious ³	RR 1.13 (0.89, 1.43)	Very low
SAE (Pramipexole- Rasagiline)	1 (Viallet)	Serious ¹	N/A	Serious ²	Serious ³	RR 0.95 (0.06, 14.75)	Very low
AE discontinuation (Pramipexole- Rasagiline)	1 (Viallet)	Serious ¹	N/A	Serious ²	Serious ³	RR 2.83 (0.79, 10.06)	Very low

	No. of						
Outcome	studies	Risk of bias	Inconsistency	Indirectness	Imprecision	Estimate (CI)	Overall quality

¹Selection of adverse events to report unclear; ²Population not treatment-naïve; ³Non-significant result

Levodopa versus dopamine agonists versus monoamine oxidase inhibitors

Outcome	No. of studies	Risk of bias	Inconsistency	Indirectness	Imprecision	Estimate (CI)	Overall quality		
Levodopa versus levo	Levodopa versus levodopa sparing (dopamine agonists and MAOBs)								
Mobility*	1 (PD MED)	Not serious	N/A	Not serious	Serious ¹	MD 1.8 [0.5, 3.0]	Moderate		
Activities of daily living*	1 (PD MED)	Not serious	N/A	Not serious	Serious ¹	MD 1.9 [0.7, 3.0]	Moderate		
Emotional wellbeing*	1 (PD MED)	Not serious	N/A	Not serious	Serious ²	MD -0.2 [-1.1, 0.7]	Moderate		
Stigma*	1 (PD MED)	Not serious	N/A	Not serious	Serious ¹	MD 1.3 [0.2, 2.3]	Moderate		
Social support*	1 (PD MED)	Not serious	N/A	Not serious	Serious ²	MD 0.1 [-0.6, 0.8]	Moderate		
Cognition*	1 (PD MED)	Not serious	N/A	Not serious	Serious ²	MD 1.0 [0.0, 2.0]	Moderate		
Communication*	1 (PD MED)	Not serious	N/A	Not serious	Serious ²	MD 0.9 [0.0, 1.8]	Moderate		
Bodily discomfort*	1 (PD MED)	Not serious	N/A	Not serious	Serious ¹	MD 1.4 [0.3, 2.4]	Moderate		
PDQ summary index	1 (PD MED)	Not serious	N/A	Not serious	Not serious	MD 1.0 [0.3, 1.7]	High		
EQ-5D utility	1 (PD MED)	Not serious	N/A	Not serious	Not serious	MD 0.03 [0.01, 0.05]	High		
Dyskinesia	1 (PD MED)	Not serious	N/A	Not serious	Not serious	HR 1.52 [1.16, 2.00]	High		

Outcome	No. of studies	Risk of bias	Inconsistency	Indirectness	Imprecision	Estimate (CI)	Overall quality
Discontinuation due to adverse events	1 (PD MED)	Not serious	N/A	Not serious	Not serious	RR 0.08 [0.04, 0.15]	High
Discontinuation due to lack of efficacy	1 (PD MED)	Not serious	N/A	Not serious	Not serious	RR 0.09 [0.04, 0.22]	High
Dopamine agonists ve	ersus MAOBs						
Mobility*	1 (PD MED)	Not serious	N/A	Not serious	Serious ²	MD 1.4 [0.0, 2.9]	Moderate
Activities of daily living*	1 (PD MED)	Not serious	N/A	Not serious	Serious ²	MD 0.3 [-1.1, 1.7]	Moderate
Emotional wellbeing*	1 (PD MED)	Not serious	N/A	Not serious	Serious ²	MD 0.3 [-0.8, 1.4]	Moderate
Stigma*	1 (PD MED)	Not serious	N/A	Not serious	Serious ²	MD 1.3 [0.0, 2.5]	Moderate
Social support*	1 (PD MED)	Not serious	N/A	Not serious	Serious ²	MD 0.8 [-0.1, 1.7]	Moderate
Cognition*	1 (PD MED)	Not serious	N/A	Not serious	Serious ¹	MD 1.7 [0.5, 2.9]	Moderate
Communication*	1 (PD MED)	Not serious	N/A	Not serious	Serious ²	MD 0.5 [-0.6, 1.5]	Moderate
Bodily discomfort*	1 (PD MED)	Not serious	N/A	Not serious	Serious ²	MD 0.7 [-0.6, 2.0]	Moderate
PDQ summary index	1 (PD MED)	Not serious	N/A	Not serious	Serious ²	MD 0.8 [0.0, 1.7]	Moderate
EQ-5D utility	1 (PD MED)	Not serious	N/A	Not serious	Serious ²	MD 0.004 [-0.01, 0.02]	Moderate

*PDQ subscale ¹Significant result but mean difference below trials defined MID ²Non-significant result

Network meta-analyses

UPDRS II (ADL): <6 months follow-up

Quality assessment					
Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Quality
Change in UPDRS II (ADL) score					
10 Mally et al., 1995; Caraceni et al., 2001; Hauser et al., 2010; Jankovic et al., 2007; Mizuno et al., 2013; Hubble et al., 1995; Palhågen et al., 1998; Parkinson Study Group 1997; Parkinson Study Group 2002; Zhang et al., 2016	Not serious	Serious ¹	Serious ²	Not serious	Low
¹ Considerable between study heterogeneity (i ² > ² Population not treatment-naïve	40%)				

UPDRS II (ADL): 6 months to 2.5 years follow-up

Quality assessment								
Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Quality			
Change in UPDRS II (ADL) score	Change in UPDRS II (ADL) score							
6 Fahn et al., 1995; Schapira et al., 2013; Palhågen et al., 1998; Poewe et al., 2011; Pahwa et al., 2014; Parkinson Study Group 2002	Not serious	Serious ¹	Serious ²	Not serious	Low			
¹ Considerable between study heterogeneity (i ² >40%) ² Population not treatment-naïve								

UPDRS III (motor): <6 months follow-up

Quality assessment					
Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Quality
Change in UPDRS III (motor) score					
10 Mally et al., 1995; Caraceni et al., 2001; Hauser et al., 2010; Jankovic et al., 2007;	Not serious	Serious ¹	Serious ²	Not serious	Low

Quality assessment						
Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Quality	
Mizuno et al., 2013; Hubble et al., 1995; Palhågen et al., 1998; Parkinson Study Group 1997; Parkinson Study Group 2002; Rascol et al., 2000						
¹ Considerable between study heterogeneity (i ² > ² Population not treatment-naïve	40%)					

UPDRS III (motor): 6 months to 2.5 years follow-up

Quality assessment							
Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Quality		
Change in UPDRS III (motor) score							
7 Fahn et al., 1995; Schapira et al., 2013; Palhågen et al., 1998; Poewe et al., 2011; Pahwa et al., 2014; Parkinson Study Group 2002; Whone et al., 2003	Not serious	Serious ¹	Serious ²	Not serious	Low		
¹ Considerable between study heterogeneity (i ² >40%) ² Population not treatment-naïve							

UPDRS total: <6 months follow-up

Quality assessment					
Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Quality
Change in UPDRS total score					
5 Adler et al., 1997; Mally et al., 1995; Hubble et al., 1995; Palhågen et al., 1998; Parkinson Study Group 1997	Not serious	Not serious	Serious ¹	Not serious	Moderate
² Population not treatment-naïve					

UPDRS total: 6 months to 2.5 years follow-up

Quality assessment					
Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Quality
Change in UPDRS total score					

Quality assessment					
Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Quality
5 Fahn et al., 1995; Schapira et al., 2013; Palhågen et al., 1998; Parkinson Study Group 2002; Olanow et al., 2009	Not serious	Not serious	Serious ²	Not serious	Moderate
² Population not treatment-naïve					

Epworth Sleep Scale

Quality assessment					
Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Quality
Change in ESS score					
3 Hauser et al., 2010; Jankovic et al., 2007; Viallet et al., 2013	Not serious	Serious ¹	Serious ²	Not serious	Low
² Population not treatment-naïve					

Meta-analyses

Treatment-naïve population

Direct meta-analysis - change in UPDRS (total) from baseline to 36 weeks/12 months (MAOBs vs placebo)



tau² (estimated amount of total heterogeneity): 0 (SE = 3.3554)

tau (square root of estimated tau² value): 0

I^2 (total heterogeneity / total variability): 0.00%

H² (total variability / sampling variability): 1.00

Test for Heterogeneity:

Q(df = 1) = 0.0497, p-val = 0.8236

Network meta-analysis - UPDRS (total) - FE model



Differences between treatments – mean and 95% confidence interval

	Treatment A					
Treatment B		Placebo	MAOB	Pramipexole	Levodopa (150/300mg)	Levodopa (600mg)
	Placebo	N/A				
	MAOB	3.07 (2.37, 3.78)	N/A			
	Pramipexole	4.80 (3.14, 6.46)	1.73 (-0.08, 3.53)	N/A		
	Levodopa (150/300mg)	5.00 (2.25, 7.76)	1.93 (-0.92, 4.77)	0.20 (-3.02, 3.42)	N/A	
	Levodopa (600mg)	8.00 (4.75, 11.25)	4.93 (1.60, 8.26)	3.20 (-0.45, 6.85)	3.00 (0.49, 5.51)	N/A

Quantifying heterogeneity/inconsistency:

tau^2 < 0.0001; l^2 = 0%

Test of heterogeneity/inconsistency:

Q d.f. p.value

0.05 1 0.8236

Network graph:



Network meta-analysis - UPDRS 2 (ADL) - FE model



Differences between treatments - mean and 95% confidence interval

	Treatment A								
Treatment B		Placebo	Selegiline	Pramipexole	Levodopa (150/300mg)	Levodopa (600mg)			
	Placebo	N/A							
	Selegiline	0.30 (-0.90, 1.50)	N/A						
	Pramipexole	1.10 (0.55, 1.65)	0.80 (-0.52, 2.12)	N/A					
	Levodopa (150/300mg)	1.60 (0.56, 2.64)	1.30 (-0.29, 2.89)	0.50 (-0.68, 1.68)	N/A				
	Levodopa (600mg)	2.20 (0.99, 3.41)	1.90 (0.20, 3.60)	1.10 (-0.23, 2.43)	0.60 (-0.29, 1.49)	N/A			

Quantifying heterogeneity/inconsistency:

tau² < 0.0001; l² = 100%

Test of heterogeneity/inconsistency:

Q d.f. p.value

0 0 <0.0001

Network graph:



Network meta-analysis - UPDRS 3 (motor) – FE model



Differences between treatments – mean and 95% confidence interval

	Treatment A								
Treatment B		Placebo	Selegiline	Levodopa (150/300mg)	Pramipexole	Levodopa (600mg)			
	Placebo	N/A							
	Selegiline	1.90 (-1.46, 5.26)	N/A						
	Levodopa (150/300mg)	2.90 (0.86, 4.94)	1.00 (-2.92, 4.93)	N/A					
	Pramipexole	3.30 (1.91, 4.69)	1.40 (-2.23, 5.03)	0.40 (-2.07, 2.86)	N/A				
	Levodopa (600mg)	5.40 (2.95, 7.85)	3.50 (-0.65, 7.65)	2.50 (0.55, 4.45)	2.10 (-0.71, 4.91)	N/A			

Quantifying heterogeneity/inconsistency:

tau² < 0.0001; l² = 100%

Test of heterogeneity/inconsistency:

Q d.f. p.value

0 0 <0.0001

Network graph:



Full population

Direct meta-analysis – short-term (≤6 months) change in UPDRS (ADL) (dopamine agonists vs placebo)



Random-Effects Model (k = 5; tau² estimator: REML)

tau² (estimated amount of total heterogeneity): 0.0000 (SE = 0.1561)

tau (square root of estimated tau² value): 0.0001

I^2 (total heterogeneity / total variability): 0.00%

H² (total variability / sampling variability): 1.00

Test for Heterogeneity:

Q(df = 4) = 5.9902, p-val = 0.3072
Direct meta-analysis – short-term (≤6 months) change in UPDRS (motor) (dopamine agonists vs placebo)



Random-Effects Model (k = 6; tau² estimator: REML)

 tau^{2} (estimated amount of total heterogeneity): 0 (SE = 0.7433)

tau (square root of estimated tau² value): 0

I^2 (total heterogeneity / total variability): 0.00%

H² (total variability / sampling variability): 1.00

Test for Heterogeneity:

Q(df = 5) = 2.2088, p-val = 0.8196

Direct meta-analysis – short-term (≤6 months) change in UPDRS (ADL) (MAOBs vs placebo)



Random-Effects Model (k = 3; tau^A2 estimator: REML)

tau² (estimated amount of total heterogeneity): 0.0000 (SE = 0.2004)

tau (square root of estimated tau² value): 0.0012

I^2 (total heterogeneity / total variability): 0.00%

H² (total variability / sampling variability): 1.00

Test for Heterogeneity:

Q(df = 2) = 4.7529, p-val = 0.0929

Direct meta-analysis – short-term (≤6 months) change in UPDRS (motor) (MAOBs vs placebo)



Random-Effects Model (k = 3; tau² estimator: REML)

tau² (estimated amount of total heterogeneity): 6.3590 (SE = 7.7656)

tau (square root of estimated tau² value): 2.5217

I² (total heterogeneity / total variability): 87.34%

H² (total variability / sampling variability): 7.90

Test for Heterogeneity:

Q(df = 2) = 10.8437, p-val = 0.0044

Direct meta-analysis – medium term (6 months – 2.5 years) change in UPDRS (total) (MAOBs vs placebo)



Random-Effects Model (k = 2; tau[^]2 estimator: REML)

- tau^A2 (estimated amount of total heterogeneity): 0 (SE = 3.3554)
- tau (square root of estimated tau² value): 0
- I^2 (total heterogeneity / total variability): 0.00%
- H² (total variability / sampling variability): 1.00

Test for Heterogeneity:

Q(df = 1) = 0.0497, p-val = 0.8236

Network meta-analysis - UPDRS 2 (ADL) - short - RE model



Differences between treatments – mean and 95% confidence interval

	Treatment A								
Treatment B		Placebo	MAOB	Dopamine agonists	Levodopa				
	Placebo	N/A							
	MAOB	1.06 (0.63, 1.49)	N/A						
	Dopamine agonists	1.10 (0.76, 1.44)	0.04 (-0.51, 0.58)	N/A					
	Levodopa	2.16 (1.46, 2.86)	1.10 (0.55, 1.65)	1.06 (0.29, 1.84)	N/A				

Quantifying heterogeneity/inconsistency:

tau^2 = 0.0743; l^2 = 54.9%

Test of heterogeneity/inconsistency:

- Q d.f. p.value
- 13.3 6 0.0385

Network graph:



Network meta-analysis - UPDRS 2 (ADL) - medium - RE model



Differences between treatments – mean and 95% confidence interval

	Placebo	MAOB	Dopamine agonists	Levodopa (low)	Levodopa (high)	Levodopa (ER)
Placebo	N/A					
MAOB	0.30 (-1.34, 1.94)	N/A				
Dopamine agonists	1.54 (0.76, 2.31)	1.54 (0.32, 2.77)	N/A			
Levodopa (low)	1.76 (0.40, 3.12)	1.70 (0.26, 3.14)	0.22 (-1.20, 1.64)	N/A		
Levodopa (high)	2.49 (1.40, 3.57)	2.57 (1.29, 3.85)	0.95 (-0.04, 1.94)	0.73 (-0.58, 2.04)	N/A	
Levodopa (ER)	3.47 (2.16, 4.79)	3.17 (1.78, 4.57)	1.94 (0.41, 3.47)	1.72 (-0.18, 3.61)	0.99 (-0.72, 2.69)	N/A

Quantifying heterogeneity/inconsistency:

tau² = 0.3201; l² = 80.9%

Test of heterogeneity/inconsistency:

Q d.f. p.value

10.47 2 0.0053

Network graph:



Network meta-analysis - UPDRS 3 (motor) – short – RE model



Differences between treatments – mean and 95% confidence interval

	Treatment A								
Treatment B		Placebo	MAOB	Dopamine agonists	Levodopa				
	Placebo	N/A							
	Dopamine agonists	2.97 (1.63, 4.31)	N/A						
	MAOB	3.40 (1.97, 4.83)	0.43 (-1.34, 2.20)	N/A					
	Levodopa	4.90 (3.00, 6.80)	1.93 (0.07, 3.79)	1.50 (-0.23, 3.23)	N/A				

Quantifying heterogeneity/inconsistency:

tau^2 = 1.0095; l^2 = 55.2%

Test of heterogeneity/inconsistency:

- Q d.f. p.value
- 15.6 7 0.0289

Network graph:



Network meta-analysis - UPDRS 3 (motor) - medium - RE model



Differences between treatments – mean and 95% confidence interval

	Placebo	МАОВ	Dopamine agonists	Levodopa (low)	Levodopa (high)	Levodopa (ER)
Placebo	N/A					
MAOB	1.90 (-2.35, 6.15)	N/A				
Dopamine agonists	3.62 (1.75, 5.50)	1.72 (-2.92, 6.37)	N/A			
Levodopa (low)	3.85 (0.83, 6.88)	1.95 (-3.26, 7.17)	0.23 (-2.99, 3.45)	N/A		
Levodopa (high)	7.25 (4.79, 9.71)	5.35 (0.44, 10.26)	3.63 (1.38, 5.88)	3.40 (0.40, 6.40)	N/A	
Levodopa (ER)	9.23 (5.60, 12.85)	7.33 (1.74, 12.91)	5.60 (1.52, 9.68)	5.37 (0.65, 10.10)	1.98 (-2.41, 6.36)	N/A

Quantifying heterogeneity/inconsistency:

tau² = 1.7971; l² = 67.0%

Test of heterogeneity/inconsistency:

Q d.f. p.value

9.09 3 0.0282

Network graph:



Network meta-analysis - UPDRS (total) – short – FE model



Differences between treatments – mean and 95% confidence interval

	Treatment A						
Treatment B		Placebo	MAOB	Dopamine agonists			
	Placebo	N/A					
	MAOB	4.02 (2.27, 5.77)	N/A				
	Dopamine agonists	4.75 (2.71, 6.80)	0.74	N/A			
			(-1.96, 3.43)				

Quantifying heterogeneity/inconsistency:

tau² = 0.0732; l² = 1.8%

Test of heterogeneity/inconsistency:

Q d.f. p.value

3.06 3 0.383

Network graph:



Network meta-analysis - UPDRS (total) – medium – FE model



Differences between treatments – mean and 95% confidence interval

	Placebo	MAOB	Dopamine agonists	Levodopa (low)	Levodopa (high)
Placebo	N/A				
MAOB	3.07 (2.37, 3.78)	N/A			
Dopamine agonists	4.64 (3.07, 6.22)	1.57 (-0.15, 3.29)	N/A		
Levodopa (low)	5.34 (2.84, 7.84)	2.26 (-0.33, 4.86)	0.69 (-2.04, 3.43)	N/A	
Levodopa (high)	8.60 (6.08, 11.12)	5.53 (2.91, 8.14)	3.96 (1.39, 6.53)	3.26 (0.92, 5.51)	N/A

Quantifying heterogeneity/inconsistency:

tau^2 < 0.0001; l^2 = 0%

Test of heterogeneity/inconsistency:

- Q d.f. p.value
- 0.38 2 0.8283

Network graph:



Network meta-analysis - Epworth sleep scale – RE model



Differences between treatments – mean and 95% confidence interval

	Treatment A							
Treatment B		Placebo	MAOB	Dopamine agonists				
	Placebo	N/A						
	MAOB	0.42	N/A					
		(-1.18, 2.03)						
	Dopamine agonists	-1.50	-1.92	N/A				
		(-2.34, -0.65)	(-2.64, -1.20)					

Quantifying heterogeneity/inconsistency:

tau^2 = 0.3508; I^2 = 94.4%

Test of heterogeneity/inconsistency:

Q d.f. p.value

17.81 1 <0.0001

Network graph:



E.2.2 Adjuvant treatment of motor symptoms

Efficacy outcomes by drug classes - Pairwise meta-analyses

Dopamine agonists vs. placebo

Outcome	No. of studies	Risk of bias	Inconsistency	Indirectness	Imprecision	Estimate (CI)	Overall quality
Off time	19 ^a	Not serious	Serious ²	Serious ³	Not serious	MD -1.42 [-1.83, -1.01]	Low
UPDRS II (ADL)	14 ^b	Not serious	Serious ²	Serious ³	Not serious ⁷	MD -1.72 [-2.16, -1.27]	Low
UPDRS III (motor)	15 ^c	Not serious	Serious ²	Serious ³	Serious ⁵	MD -4.09 [-5.25, -2.92]	Very low
PDQ-39	2 ^d	Serious ¹	Serious ²	Serious ³	Serious ⁶	MD -1.88 [-5.40, 1.64]	Very Low
PDQUALIF	1 ^e	Serious ¹	N/A	Serious ³	Serious ⁴	MD -3.22 [-6.86, 0.42]	Very Low

a Stowe Cochrane review 2010 (n=15: Interntl; Germany; Spain; UK; USA I; N America; Aust/Germ; CLEOPATRA; Denmark; Europe; US/Canada; EASE-PD; France/Eng; UK/Israel; USA) Nicholas 2014; Nomoto 2014; Pahwa 2007; Poewe 2007

b Stowe Cochrane review 2010 (n=6: Spain; USA I; Aust/Germ; CLEOPATRA; Europe; H Kong/Taiw); Mizuno 2003; Mizuno 2007; Nicholas 2014; Nomoto 2014; Pahwa 2007; Poewe 2007; PSG 2007; Watts 2010

c Stowe Cochrane review 2010 (n=7: Spain; USA I; Aust/Germ; CLEOPATRA; Europe; H Kong/Taiw; EASE-PD); Mizuno 2003; Mizuno 2007; Nicholas 2014; Nomoto 2014; Pahwa 2007; Poewe 2007; PSG 2007; Watts 2010

d Poewe 2007; Watts 2010

e PSG 2007

¹Individual study(ies) at risk of bias; ²Considerable between study heterogeneity (i² >40%); ³Population not as defined in protocol; ⁴Non-significant result; ⁵Cl cross MID: between 3.25 (Horváth et al., 2015) and 5 points (Schrag et al., 2006); ⁶Cl cross MID of 1.6 points (Peto et al., 2001); ⁷Cl do not cross MID of 3 points (Schrag et al., 2006)

COMTIs versus placebo

Outcome	No. of studies	Risk of bias	Inconsistency	Indirectness	Imprecision	Estimate (CI)	Overall quality
Off time	13 ^a	Not serious	Not serious	Serious ³	Not serious	MD -0.81 [-1.01, <i>-</i> 0.60]	Moderate
UPDRS II (ADL)	12 ^b	Not serious	Not serious	Serious ³	Not serious⁵	MD -0.99 [-1.35, -0.63]	Moderate
UPDRS III (motor)	13°	Not serious	Not serious	Serious ³	Not serious ⁶	MD -2.11 [-2.74, -1.47]	Moderate
PDQ-39	1 ^d	Not serious	N/A	Serious ³	Serious ⁴	MD 6.90 [-4.05, 17.85]	Low

a Stowe Cochrane review 2010 (n=12: Celomen; ComQol; INT-01; LARGO; Nomecomt; Sth Korea; UK/Irish; China; Europe; TFSG I; TFSG 3; TIPS I); Fenelon 2003

b Stowe Cochrane review 2010 (n=10: Celomen; ComQol; INT-02; Nomecomt; Sth Korea; UK/Irish; TFSG 3; TIPS I; TIPS II; US/Canada); Fenelon 2003; Tolosa 2014

c Stowe Cochrane review 2010 (n=12: Celomen; ComQol; Interntl; LARGO; Nomecomt; Sth Korea; UK/Irish; Europe; TFSG 3; TIPS I; TIPS II; US/Canada); Tolosa 2014

d Tolosa 2014

¹Individual study(ies) at risk of bias; ²Considerable between study heterogeneity (i² >40%); ³Population not as defined in protocol; ⁴CI cross MID of 1.6 points (Peto et al., 2001); ⁵CI do not cross MID of 3 points (Schrag et al., 2006); ⁶CI do not cross MID: between 3.25 (Horváth et al., 2015) and 5 points (Schrag et al., 2006); ⁶CI do not cross MID: between 3.25 (Horváth et al., 2015) and 5 points (Schrag et al., 2006); ⁶CI do not cross MID of 3 points (Schrag et al., 2006); ⁶CI do not cross MID: between 3.25 (Horváth et al., 2015) and 5 points (Schrag et al., 2006); ⁶CI do not cross MID: between 3.25 (Horváth et al., 2015) and 5 points (Schrag et al., 2006); ⁶CI do not cross MID: between 3.25 (Horváth et al., 2015) and 5 points (Schrag et al., 2006); ⁶CI do not cross MID: between 3.25 (Horváth et al., 2015) and 5 points (Schrag et al., 2006); ⁶CI do not cross MID: between 3.25 (Horváth et al., 2015) and 5 points (Schrag et al., 2006); ⁶CI do not cross MID: between 3.25 (Horváth et al., 2015) and 5 points (Schrag et al., 2006); ⁶CI do not cross MID: between 3.25 (Horváth et al., 2015) and 5 points (Schrag et al., 2006); ⁶CI do not cross MID: between 3.25 (Horváth et al., 2015) and 5 points (Schrag et al., 2006); ⁶CI do not cross MID: between 3.25 (Horváth et al., 2015) and 5 points (Schrag et al., 2006); ⁶CI do not cross MID: between 3.25 (Horváth et al., 2015) and 5 points (Schrag et al., 2006); ⁶CI do not cross MID et al., 2015); ⁶CI do not cro

MAOBIs versus placebo

Outcome	No. of studies	Risk of bias	Inconsistency	Indirectness	Imprecision	Estimate (CI)	Overall quality
Off time	4 ^a	Not serious	Not serious	Serious ¹	Not serious	MD -0.98 [-1.22, -0.74]	Moderate
UPDRS II (ADL)	1 ^b	Not serious	Not serious	Serious ¹	Not serious ²	MD -1.85 [-2.62, -1.08]	Moderate
UPDRS III (motor)	2 ^c	Not serious	N/A	Serious ¹	Not serious ³	MD -2.29 [-3.05, -1.54]	Moderate

Outcome	No. of studies	Risk of bias	Inconsistency	Indirectness	Imprecision	Estimate (CI)	Overall quality	
a Stowe Cochrane review 2010 (n=3: LARGO; PRESTO; USA); Zhang 2013								
b Zhang 2013	b Zhang 2013							
c Stowe 2010 (n=1: LA	ARGO); Zhang	j 2013						
1				() 0000 20				

¹Population not as defined in protocol; ²CI do not cross MID of 3 points (Schrag et al., 2006); ³CI do not cross MID: between 3.25 (Horváth et al., 2015) and 5 points (Schrag et al., 2006)

Dopamine agonists versus COMTIs

Outcome	No. of studies	Risk of bias	Inconsistency	Indirectness	Imprecision	Estimate (CI)	Overall quality
UPDRS II (ADL)	2ª	Serious ¹	Not serious	Serious ²	Not serious ⁴	MD 0.40 [-0.48, 1.27]	Low
UPDRS III (motor)	2ª	Serious ¹	Not serious	Serious ²	Not serious ⁵	MD -0.10 [-2.06, 1.86]	Low
Off time	2 ^a	Serious ¹	Not serious	Serious ²	Serious ³	MD -0.11 [-0.83, 0.60]	Very Low
PDQ-39	1 ^b	Serious ¹	N/A	Serious ²	Serious ⁶	MD -2.90 [-6.38, 0.58]	Very low

a Deane 2004 (n=1); Deuschl 2007

b Deuschl 2007

¹Individual study(ies) at risk of bias; ²Population not as defined in protocol; ³Non-significant result; ⁴CI do not cross MID of 3 points (Schrag et al., 2006); ⁵CI do not cross MID: between 3.25 (Horváth et al., 2015) and 5 points (Schrag et al., 2006); ⁶CI cross MID of 1.6 points (Peto et al., 2001)

Amantadine versus placebo

Outcome	No. of studies	Risk of bias	Inconsistency	Indirectness	Imprecision	Estimate (CI)	Overall quality
Hyperkinesia (CDRS)	1 ^a	Not serious	N/A	Serious ³	Serious ⁴	MD -6.20 [-14.37, 1.97]	Low
Dystonia (CDRS)	1 ^a	Not serious	N/A	Serious ³	Serious ⁴	MD -0.40 [-4.06, 3.26]	Low
UPDRS II	1 ^a	Not serious	N/A	Serious ³	Serious ⁵	MD -1.70	Low

Outcome	No. of studies	Risk of bias	Inconsistency	Indirectness	Imprecision	Estimate (CI)	Overall quality
						[-9.05, 5.65]	
UPDRS III	1 ^a	Not serious	N/A	Serious ³	Serious ⁶	MD -2.40 [-9.39, 4.59]	Low

a da Silvia-Junior 2005

¹Individual study(ies) at risk of bias; ²Considerable between study heterogeneity ($i^2 > 40\%$); ³Population not as defined in protocol; 4Non-significant result; 5Cl cross MID of 3 points (Schrag et al., 2006); 6Cl cross MID: between 3.25 (Horváth et al., 2015) and 5 points (Schrag et al., 2006)

Safety outcomes by individual drugs - Pairwise meta-analyses

Ropinirole versus placebo

Outcome	No. of studies	Risk of bias	Inconsistency	Indirectness	Imprecision	Estimate (CI)	Overall quality
Dyskinesia	7 ^a	Serious ¹	Serious ²	Serious ³	Serious ⁴	RR 2.36 [0.77, 7.22]	Very Low
Hallucinations	3 ^b	Not serious	Not serious	Serious ³	Not serious	RR 5.97 [2.23, 16.02]	Moderate
Mortality	3°	Not serious	Not serious	Serious ³	Serious ⁴	RR 0.29 [0.03, 2.77]	Low
Any AEs	7 ^d	Serious ¹	Not serious	Serious ³	Not serious	RR 1.15 [1.08, 1.23]	Low
SAEs	3 ^e	Serious ¹	Not serious	Serious ³	Serious ⁴	RR 0.94 [0.56, 1.57]	Very Low
AE discontinuation	7 ^f	Not serious	Not serious	Serious ³	Serious ⁴	RR 1.11 [0.80, 1.53]	Low
Psychosis (Parkinson's Psychosis Rating Scale)	1 ⁹	Not serious	N/A	Serious ³	Serious ⁴	MD 0.30 [-0.20, 0.80]	Low

a Stowe Cochrane review 2010 (n=3: EASE-PD; France/Eng; USA); Lieberman 1997; Mizuno 2010; Mizuno 2014; Watts 2010

b Stowe Cochrane review 2010 (n=1: EASE-PD); Mizuno 2010; Mizuno 2014

c Stowe Cochrane review 2010 (n=3: EASE-PD; France/Eng; UK/Israel)

Outcome	No. of studies	Risk of bias	Inconsistency	Indirectness	Imprecision	Estimate (CI)	Overall quality	
d Stowe Cochrane review 2010 (n=3: EASE-PD; France/Eng; UK/Israel); Mizuno 2010; Mizuno 2014; Pahwa 2007; Watts 2010								
e Mizuno 2010; Mizuno 2014; Watts 2010								
f Stowe Cochrane revi	ew 2010 (n=4:	EASE-PD; France	/Eng; UK/Israel; USA)); Mizuno 2010; Miz	zuno 2014; Watts 2	010		
g Watts 2010								
¹ Individual study(ies) a	at risk of bias; ²	Considerable betw	een study heterogene	eity; ³ Population not	as defined in proto	col; ⁴ Non-significant re	esult	

Rotigotine versus placebo

Outcome	No. of studies	Risk of bias	Inconsistency	Indirectness	Imprecision	Estimate (CI)	Overall quality
Dyskinesia	5 ^a	Not serious	Not serious	Serious ³	Not serious	RR 3.06 [1.95, 4.81]	Moderate
Hallucinations	5 ^a	Not serious	Not serious	Serious ³	Not serious	RR 3.89 [1.82, 8.30]	Moderate
Any AEs	4 ^b	Not serious	Serious ²	Serious ³	Serious ⁴	RR 1.09 [0.99, 1.20]	Low
SAEs	3°	Not serious	Not serious	Serious ³	Serious ⁴	RR 0.61 [0.31, 1.19]	Low
AE discontinuation	5 ^a	Serious ¹	Not serious	Serious ³	Serious ⁴	RR 0.87 [0.63, 1.21]	Very Low
Mortality	1 ^d	Not serious	N/A	Serious ³	Serious ⁴	RR 1.34 [0.06, 27.69]	Low
Impulse Control Disorder	1 ^d	Not serious	N/A	Serious ³	Serious ⁴	RR 2.93 [0.16, 52.61]	Low

a Lewitt 2007; Mizuno 2014; Nicholas 2014; Nomoto 2014; Poewe 2007

b Mizuno 2014; Nicholas 2014; Nomoto 2014; Poewe 2007

c Mizuno 2014; Nicholas 2014; Nomoto 2014

d Nicholas 2014

¹Individual study(ies) at risk of bias; ²Considerable between study heterogeneity (i² >40%); ³Population not as defined in protocol; ⁴Non-significant result

Pramipexole versus placebo

Outcome	No. of studies	Risk of bias	Inconsistency	Indirectness	Imprecision	Estimate (CI)	Overall quality
Dyskinesia	10 ^a	Not serious	Not serious	Serious ³	Not serious	RR 1.92 [1.61, 2.29]	Moderate
Hallucinations	9 ^b	Not serious	Not serious	Serious ³	Not serious	RR 2.86 [1.99, 4.09]	Moderate
Any AEs	8 ^c	Not serious	Not serious	Serious ³	Not serious	RR 1.08 [1.01, 1.14]	Moderate
SAEs	3 ^d	Serious ¹	Not serious	Serious ³	Serious ⁴	1.49 [0.64, 3.44]	Very Low
AE discontinuation	8°	Not serious	Not serious	Serious ³	Serious ⁴	RR 0.86 [0.66, 1.12]	Low

a Stowe Cochrane review 2010 (n=6: Aust/Germ; CLEOPATRA; Denmark; Europe; Interntl; US/Canada); Mizuno 2003; Poewe 2007; PSG 2007; Schapira 2011

b Stowe Cochrane review 2010 (n=5: Aust/Germ; CLEOPATRA; Europe; Interntl; US/Canada); Mizuno 2003; Poewe 2007; PSG 2007; Schapira 2011

c Stowe Cochrane review 2010 (n=5: Aust/Germ; CLEOPATRA; Denmark; Interntl; US/Canada); Mizuno 2003; Poewe 2007; Schapira 2011

d Mizuno 2003; PSG 2007; Schapira 2011

¹Individual study(ies) at risk of bias; ²Considerable between study heterogeneity (i² >40%); ³Population not as defined in protocol; ⁴Non-significant result

Cabergoline versus placebo

Outcome	No. of studies	Risk of bias	Inconsistency	Indirectness	Imprecision	Estimate (CI)	Overall quality
Dyskinesia	3 ^a	Not serious	Not serious	Serious ³	Not serious	RR 1.29 [1.01, 1.64]	Moderate
Hallucinations	3ª	Not serious	Not serious	Serious ³	Serious ⁴	RR 2.18 [0.74, 6.46]	Low
Mortality	1 ^b	Not serious	N/A	Serious ³	Serious ⁴	RR 0.33 [0.01, 7.72]	Low
Any AEs	3 ^a	Not serious	Not serious	Serious ³	Not serious	RR 1.17 [1.03, 1.34]	Moderate
AE discontinuation	3 ^a	Not serious	Serious ²	Serious ³	Serious ⁴	RR 1.25	Very Low

Outcome	No. of studies	Risk of bias	Inconsistency	Indirectness	Imprecision	Estimate (CI)	Overall quality
						[0.48, 3.22]	
a Stowe Cochrane rev	view 2010 (n=3:	Spain; USA I; US	A 2)				

b Stowe Cochrane review 2010 (n=1: Spain)

¹Individual study(ies) at risk of bias; ²Considerable between study heterogeneity (i² >40%); ³Population not as defined in protocol; ⁴Non-significant result

Bromocriptine versus placebo

Outcome	No. of studies	Risk of bias	Inconsistency	Indirectness	Imprecision	Estimate (CI)	Overall quality
Dyskinesia	3 ^a	Not serious	Not serious	Serious ³	Not serious	RR 1.82 [1.20, 2.76]	Moderate
Hallucinations	3ª	Not serious	Serious ²	Serious ³	Serious ⁴	RR 1.93 [0.49, 7.56]	Low
Any AEs	3ª	Not serious	Not serious	Serious ³	Not serious	RR 1.17 [1.03, 1.34]	Moderate
AE discontinuation	5 ^b	Not serious	Not serious	Serious ³	Serious ⁴	RR 1.02 [0.71, 1.47]	Low

a Stowe Cochrane review 2010 (n=2: Interntl; Japan); Mizuno 2003

b Stowe Cochrane review 2010 (n=4: Interntl; Japan; Rotterdam; South Africa); Mizuno 2003

¹Individual study(ies) at risk of bias; ²Considerable between study heterogeneity (i² >40%); ³Population not as defined in protocol; ⁴Non-significant result

Pergolide versus placebo

Outcome	No. of studies	Risk of bias	Inconsistency	Indirectness	Imprecision	Estimate (CI)	Overall quality
Dyskinesia	1 ^a	Not serious	N/A	Serious ³	Not serious	RR 2.54 [1.93, 3.34]	Moderate
Hallucinations	1 ^a	Not serious	N/A	Serious ³	Not serious	RR 4.29 [1.81, 10.18]	Moderate
Mortality	1 ^a	Not serious	N/A	Serious ³	Serious ⁴	RR 0.49 [0.05, 5.41]	Low
AE discontinuation	1 ^a	Not serious	N/A	Serious ³	Serious ⁴	RR 2.23	Low

Outcome	No. of studies	Risk of bias	Inconsistency	Indirectness	Imprecision	Estimate (CI)	Overall quality		
						0.99, 4.99]			
a Stowe Cochrane review (n=1: N America)									

¹Individual study(ies) at risk of bias; ²Considerable between study heterogeneity (i² >40%); ³Population not as defined in protocol; ⁴Non-significant result

Entacapone versus placebo

Outcome	No. of studies	Risk of bias	Inconsistency	Indirectness	Imprecision	Estimate (CI)	Overall quality
Dyskinesia	11 ^a	Not serious	Not serious	Serious ³	Not serious	RR 2.01 [1.67, 2.42]	Moderate
Hallucinations	8 ^b	Not serious	Serious ²	Serious ³	Serious ⁴	RR 0.43 [0.03, 6.84]	Very Low
Mortality	1 ^c	Not serious	N/A	Serious ³	Serious ⁴	RR 0.40 [0.09, 1.79]	Low
Any AEs	10 ^d	Serious1	Serious ²	Serious ³	Not serious	RR 1.39 [1.07, 1.81]	Very Low
SAEs	3 ^e	Not serious	Not serious	Serious ³	Serious ⁴	RR 0.91 [0.39, 2.12]	Low
AE discontinuation	12 ^f	Not serious	Not serious	Serious ³	Not serious	RR 1.51 [1.17, 1.95]	Moderate

a Stowe Cochrane review 2010 (n=10: Celomen; ComQol; Filomen; INT-02; Japan; LARGO; Nomecomt; Seesaw; Sth Korea; UK/Irish); Fenelon 2003

b Stowe Cochrane review 2010 (n=7: Celomen; INT-02; LARGO; Nomecomt; Seesaw; Sth Korea; UK/Irish); Fenelon 2003

c Stowe Cochrane review 2010 (n=1: Filomen)

d Stowe Cochrane review 2010 (n=7: Celomen; ComQol; INT-02; Japan; LARGO; Seesaw; UK/Irish;); Fenelon 2003; Destee 2009; Tolosa 2014

e Fenelon 2003; Destee 2009; Tolosa 2014

f Stowe Cochrane review 2010 (n=9: Celomen; ComQol; Filomen; INT-02; Interntl; Japan; LARGO; Nomecomt; Seesaw); Fenelon 2003; Destee 2009; Tolosa 2014

¹Individual study(ies) at risk of bias; ²Considerable between study heterogeneity (i² >40%); ³Population not as defined in protocol; ⁴Non-significant result

Tolcapone versus placebo

Outcome	No. of studies	Risk of bias	Inconsistency	Indirectness	Imprecision	Estimate (CI)	Overall quality
Dyskinesia	6 ^a	Not serious	Not serious	Serious ³	Not serious	RR 2.58 [1.93, 3.44]	Moderate
Hallucinations	4 ^b	Not serious	Serious ²	Serious ³	Not serious	RR 2.50 [1.23, 5.06]	Low
Any AEs	4 ^b	Not serious	Not serious	Serious ³	Not serious	RR 1.22 [1.10, 1.34]	Moderate
AE discontinuation	5°	Not serious	Serious ²	Serious ³	Serious ⁴	RR 1.47 [0.88, 2.46]	Very Low

a Stowe Cochrane review 2010 (n=6: China; Europe; TFSG 3; TIPS I; TIPS II; US/Canada)

b Stowe Cochrane review 2010 (n=4: TFSG 3; TIPS I; TIPS II; US/Canada)

c Stowe Cochrane review 2010 (n=5: Europe; TFSG 3; TIPS I; TIPS II; US/Canada)

¹Individual study(ies) at risk of bias; ²Considerable between study heterogeneity (i² >40%); ³Population not as defined in protocol; ⁴Non-significant result

Rasagiline versus placebo

Outcome	No. of studies	Risk of bias	Inconsistency	Indirectness	Imprecision	Estimate (CI)	Overall quality
Dyskinesia	2 ^a	Not serious	Not serious	Serious ³	Serious ⁴	RR 1.19 [0.53, 2.65]	Low
Hallucinations	1 ^b	Not serious	N/A	Serious ³	Serious ⁴	RR 1.65 [0.40, 6.83]	Low
Any AEs	3°	Not serious	Not serious	Serious ³	Serious ⁴	RR 1.06 [0.93, 1.22]	Low
SAEs	1 ^d	Not serious	N/A	Serious ³	Serious ⁴	RR 1.05 [0.07, 16.60]	Low
AE discontinuation	2 ^a	Not serious	Not serious	Serious ³	Serious ⁴	RR 0.59 [0.28, 1.28]	Low

a Stowe Cochrane review 2010 (n=1: LARGO); Zhang 2013

b Stowe Cochrane review 2010 (n=1: LARGO)

c Stowe Cochrane review 2010 (n=2: LARGO; PRESTO); Zhang 2013

Outcome	No. of studies	Risk of bias	Inconsistency	Indirectness	Imprecision	Estimate (CI)	Overall quality
d Zhang 2013							

¹Individual study(ies) at risk of bias; ²Considerable between study heterogeneity; ³Population not as defined in protocol; ⁴Non-significant result

Selegiline versus placebo

Outcome	No. of studies	Risk of bias	Inconsistency	Indirectness	Imprecision	Estimate (CI)	Overall quality
Dyskinesia	3 ^a	Serious ¹	Serious ²	Serious ³	Serious ⁴	RR 0.86 [0.44, 1.69]	Very Low
Hallucinations	1 ^b	Not serious	N/A	Serious ³	Serious ⁴	RR 2.76 [0.30, 25.60]	Low
Any AEs	3ª	Serious ¹	Not serious	Serious ³	Serious ⁴	RR 1.08 [0.88, 1.33]	Very Low
SAEs	1 ^c	Serious ¹	N/A	Serious ³	Serious ⁴	RR 4.00 [0.51, 31.10]	Very Low
AE discontinuation	3 ^a	Serious ¹	Serious ²	Serious ³	Serious ⁴	RR 1.72 [0.14, 20.91]	Very Low

a Stowe Cochrane review 2010 (n=2: Norw/Fin; USA); Ondo 2007

b Stowe Cochrane review 2010 (n=1: USA)

c Ondo 2007

¹Individual study(ies) at risk of bias; ²Considerable between study heterogeneity (i² >40%); ³Population not as defined in protocol; ⁴Non-significant result

Amantadine versus placebo

Outcome	No. of studies	Risk of bias	Inconsistency	Indirectness	Imprecision	Estimate (CI)	Overall quality
Hyperkinesia (CDRS)	1 ^a	Not serious	N/A	Serious ³	Serious ⁴	MD -6.20 [-14.37, 1.97]	Low
Dystonia (CDRS)	1 ^a	Not serious	N/A	Serious ³	Serious ⁴	MD -0.40 -4.06, 3.26]	Low

Outcome	No. of studies	Risk of bias	Inconsistency	Indirectness	Imprecision	Estimate (CI)	Overall quality
UPDRS II	1 ^a	Not serious	N/A	Serious ³	Serious⁵	MD -1.70 -9.05, 5.65]	Low
UPDRS III	1 ^a	Not serious	N/A	Serious ³	Serious ⁶	MD -2.40 [-9.39, 4.59]	Low

a da Silvia-Junior 2005

¹Individual study(ies) at risk of bias; ²Considerable between study heterogeneity (i² >40%); ³Population not as defined in protocol; ⁴Non-significant result; ⁵CI cross MID of 3 points (Schrag et al., 2006); ⁶CI cross MID: between 3.25 (Horváth et al., 2015) and 5 points (Schrag et al., 2006)

Ropinirole versus Rotigotine

Outcome	No. of studies	Risk of bias	Inconsistency	Indirectness	Imprecision	Estimate (CI)	Overall quality
Dyskinesia	1 ^a	Not serious	N/A	Serious ³	Serious ⁴	RR 0.86 [0.51, 1.43]	Low
Hallucinations	1 ^a	Not serious	N/A	Serious ³	Serious ⁴	RR 2.01 [0.51, 7.91]	Low
Any AEs	1 ^a	Not serious	N/A	Serious ³	Not serious	RR 0.88 [0.80, 0.97]	Moderate
SAEs	1 ^a	Not serious	N/A	Serious ³	Serious ⁴	RR 0.72 [0.23, 2.22]	Low
AE discontinuation	1 ^a	Not serious	N/A	Serious ³	Serious ⁴	RR 1.01 [0.48, 2.10]	Low
a Mizuno 2014							

¹Individual study(ies) at risk of bias; ²Considerable between study heterogeneity (i² >40%); ³Population not as defined in protocol; ⁴Non-significant result

Ropinirole versus Bromocriptine

Outcome	No. of studies	Risk of bias	Inconsistency	Indirectness	Imprecision	Estimate (CI)	Overall quality
Dyskinesia	2 ^a	Serious ¹	Not serious	Serious ³	Serious ⁴	RR 1.44 [0.66, 3.16]V	Very Low
Hallucinations	2 ^a	Serious ¹	Not serious	Serious ³	Serious ⁴	RR 0.76	Very Low

Outcome	No. of studies	Risk of bias	Inconsistency	Indirectness	Imprecision	Estimate (CI)	Overall quality		
						[0.27, 2.15]			
a Clarke Cochrane review 2001b (n=2)									

¹Individual study(ies) at risk of bias; ²Considerable between study heterogeneity (i² >40%); ³Population not as defined in protocol; ⁴Non-significant result

Pramipexole versus Bromocriptine

Outcome	No. of studies	Risk of bias	Inconsistency	Indirectness	Imprecision	Estimate (CI)	Overall quality
Dyskinesia	2 ^a	Not serious	Not serious	Serious ³	Not serious	RR 2.33 [1.14, 4.74]	Moderate
Hallucinations	1 ^b	Not serious	N/A	Serious ³	Serious ⁴	RR 0.90 [0.46, 1.75]	Low
Any AEs	1 ^b	Not serious	N/A	Serious ³	Serious ⁴	RR 0.94 [0.85, 1.04]	Low
SAEs	1 ^b	Not serious	N/A	Serious ³	Serious ⁴	RR 7.14 [0.37, 136.43]	Low
AE discontinuation	1 ^b	Not serious	N/A	Serious ³	Serious ⁴	RR 0.69 [0.29, 1.61]	Low

a Stowe Cochrane review 2010 (n=1: Interntl); Mizuno 2003

b Mizuno 2003

¹Individual study(ies) at risk of bias; ²Considerable between study heterogeneity (i² >40%); ³Population not as defined in protocol; ⁴Non-significant result

Rotigotine versus Pramipexole

Outcome	No. of studies	Risk of bias	Inconsistency	Indirectness	Imprecision	Estimate (CI)	Overall quality
Dyskinesia	1 ^a	Not serious	N/A	Serious ³	Serious ⁴	RR 0.76 [0.46, 1.25]	Low
Hallucinations	1 ^a	Not serious	N/A	Serious ³	Serious ⁴	RR 0.70 [0.32, 1.55]	Low
Any AEs	1 ^a	Not serious	N/A	Serious ³	Serious ⁴	RR 1.00 [0.88, 1.14]	Low

Outcome	No. of studies	Risk of bias	Inconsistency	Indirectness	Imprecision	Estimate (CI)	Overall quality
AE discontinuation	1 ^a	Not serious	N/A	Serious ³	Serious ⁴	RR 0.77 [0.36, 1.66]	Low
a Poewe 2007							

¹Individual study(ies) at risk of bias; ²Considerable between study heterogeneity (i² >40%); ³Population not as defined in protocol; ⁴Non-significant result

Pramipexole versus Pergolide

Outcome	No. of studies	Risk of bias	Inconsistency	Indirectness	Imprecision	Estimate (CI)	Overall quality
Any AEs	1 ^a	Serious ¹	N/A	Serious ³	Serious ⁴	RR 0.80 [0.52, 1.24]	Very Low
AE discontinuation	1 ^a	Serious ¹	N/A	Serious ³	Serious ⁴	RR 1.30 [0.24, 6.96]	Very Low
a Dalstanava 2002							

a Rektorova 2003

¹Individual study(ies) at risk of bias; ²Considerable between study heterogeneity (i² >40%); ³Population not as defined in protocol; ⁴Non-significant result

Cabergoline versus Bromocriptine

Outcome	No. of studies	Risk of bias	Inconsistency	Indirectness	Imprecision	Estimate (CI)	Overall quality
Dyskinesia	5 ^a	Not serious	Not serious	Serious ³	Not serious	RR 1.49 [1.04, 2.13]	Moderate
Hallucinations	5 ^a	Not serious	Not serious	Serious ³	Serious ⁴	RR 1.31 [0.89, 1.94]	Low

a Clarke Cochrane review 2001a (n=5)

¹Individual study(ies) at risk of bias; ²Considerable between study heterogeneity (i² >40%); ³Population not as defined in protocol; ⁴Non-significant result

Cabergoline versus Entacapone

Outcome	No. of studies	Risk of bias	Inconsistency	Indirectness	Imprecision	Estimate (CI)	Overall quality
Hallucinations	1 ^a	Serious ¹	N/A	Serious ³	Serious ⁴	RR 1.04 [0.22, 4.99]	Very Low

Outcome	No. of studies	Risk of bias	Inconsistency	Indirectness	Imprecision	Estimate (CI)	Overall quality
Any AEs	1 ^a	Serious ¹	N/A	Serious ³	Serious ⁴	RR 0.99 [0.74, 1.32]	Very Low
SAEs	1 ^a	Serious ¹	N/A	Serious ³	Serious ⁴	RR 0.52 [0.13, 2.00]	Very Low
AE discontinuation	1 ^a	Serious ¹	N/A	Serious ³	Serious ⁴	RR 1.63 [0.67, 4.00]	Very Low

a Deuschl 2007

¹Individual study(ies) at risk of bias; ²Considerable between study heterogeneity (i² >40%); ³Population not as defined in protocol; ⁴Non-significant result

Bromocriptine versus Tolcapone

Outcome	No. of studies	Risk of bias	Inconsistency	Indirectness	Imprecision	Estimate (CI)	Overall quality
Dyskinesia	1 ^a	Serious ¹	N/A	Serious ³	Serious ⁴	RR 0.74 [0.51, 1.06]	Very Low
Hallucinations	1 ^a	Serious ¹	N/A	Serious ³	Serious ⁴	RR 6.81 [0.86, 53.98]	Very Low
a Dean Cochrane revi	ew 2004 (n=1)						

¹Individual study(ies) at risk of bias; ²Considerable between study heterogeneity (i² >40%); ³Population not as defined in protocol; ⁴Non-significant result

Pergolide versus Tolcapone

Outcome	No. of studies	Risk of bias	Inconsistency	Indirectness	Imprecision	Estimate (CI)	Overall quality
Dyskinesia	1 ^a	Serious ¹	N/A	Serious ³	Not serious	RR 0.51 [0.34, 0.78]	Low
AE discontinuation	1 ^a	Serious ¹	N/A	Serious ³	Not serious	RR 2.97 [1.12, 7.87]	Low
a Dean Cochrane revi	ew 2004 (n=1)						

¹Individual study(ies) at risk of bias; ²Considerable between study heterogeneity ($i^2 > 40\%$); ³Population not as defined in protocol; ⁴Non-significant result

Entacapone versus Tolcapone

Outcome	No. of studies	Risk of bias	Inconsistency	Indirectness	Imprecision	Estimate (CI)	Overall quality
Dyskinesia	1 ^a	Not serious	N/A	Serious ³	Serious ⁴	RR 0.96 [0.59, 1.56]	Low
Hallucinations	1 ^a	Not serious	N/A	Serious ³	Serious ⁴	RR 7.00 [0.37, 133.22]	Low
Any AEs	1 ^a	Not serious	N/A	Serious ³	Serious ⁴	RR 0.93 [0.70, 1.24]	Low
SAEs	1 ^a	Not serious	N/A	Serious ³	Serious ⁴	RR 0.17 [0.02, 1.35]	Low
AE discontinuation	1 ^a	Not serious	N/A	Serious ³	Serious ⁴	RR 3.00 [0.12, 72.49]	Low
a ESS 2007							

¹Individual study(ies) at risk of bias; ²Considerable between study heterogeneity (i² >40%); ³Population not as defined in protocol; ⁴Non-significant result

Network meta-analyses

OFF time (hours)

Quality assessment							
Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Quality		
Change in OFF time							
35	Not serious ¹	Not serious	Serious ²	Not serious	Moderate		
DAs vs. placebo n=19							
COMTIs vs. placebo n=13							
MAOBIs vs. placebo n=3							
¹ Individual studies at risk of bias, but o	verall risk of bias rated lo	w due to consistency of effect	between studies at high a	nd low risk of bias			

²Considered serious as population is not as defined in protocol

UPDRS II (ADL)

Quality assessment					
Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Quality
Change in UPDRS II score					
30 DA vs. placebo n=14 COMTIs vs. placebo n=12 Amantadine vs. placebo n=3 DA vs. COMTIs n=3	Not serious ¹	Serious ²	Serious ³	Not serious	Low
¹ Individual studies at risk of bias, but or ² Considerable between study heteroge ³ Considered serious as population is n	verall risk of bias rated lo neity (l²>40%) ot as defined in protocol	w due to consistency of effect	between studies at high a	nd low risk of bias	

UPDRS III (motor)

Quality assessment					<u>i</u>
Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Quality
Change in UPDRS III score					
34 DAs vs. placebo n=15 COMTIs vs. placebo n=13 MAOBIs vs. placebo n=2 Amantadine vs. placebo n=1 DAs vs. COMTIs n=3	Not serious ¹	Serious ²	Serious ³	Not serious	Low
¹ Individual studies at risk of bias, but o ² Considerable between study heteroge ³ Considered serious as population is n	verall risk of bias rated lov neity (l²>40%) ot as defined in protocol	w due to consistency of effect	between studies at high a	nd low risk of bias	

PDQ-39

Quality assessment									
Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Quality				
Change in PDQ-39 score									
4	Serious ¹	Serious ²	Serious ³	Not serious	Very Low				
Quality assessment									
---	---------------------------	---------------	--------------	-------------	---------	--	--	--	--
Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Quality				
DA vs. placebo n=2									
COMTIs vs. placebo n=1									
DAs vs. COMTIs n=1									
¹ Individual studies at risk of bias									
² Considerable between study heterogeneity (I ² >40%)									
³ Considered serious as population is n	ot as defined in protocol								

Dyskinesia

Quality assessment								
Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Quality			
Dyskinesia								
65	Not serious ¹	Not serious	Serious ²	Not serious	Moderate			
DAs vs. placebo=29								
COMTIs vs. placebo n=17								
MAOBIs vs. placebo n=5								
DAs vs. DAs n=11								
DAs vs. COMTIs n=2								
COMTI vs. COMTI n=1								
¹ Individual studies at risk of bias, but or ² Considered serious as population is n	verall risk of bias rated lo ot as defined in protocol	w due to consistency of effect	between studies at high a	nd low risk of bias				

Hallucinations

Quality assessment								
Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Quality			
Hallucinations								
51 DA vs. placebo n=24 COMTIs vs. placebo n=12 MAOBIs vs. placebo =2 DA vs. DA n=10	Not serious ¹	Not serious	Serious ²	Not serious	Moderate			

Quality assessment								
Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Quality			
DA vs. COMT n=2								
COMT vs. COMT n=1								
¹ Individual studies at risk of bias, but overall risk of bias rated low due to consistency of effect between studies at high and low risk of bias								
² Considered serious as population is n	ot as defined in protocol	-	-					

Mortality

Quality assessment										
Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Quality					
Mortality										
8	Not serious	Not serious	Serious ¹	Not serious	Moderate					
DAs vs. placebo n=6										
COMTIs vs. placebo n=2										
¹ Considered serious as population is n	ot as defined in protocol									

Serious adverse events (SAEs)

Quality assessment								
Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Quality			
SAEs								
18 DAs vs. placebo n=9 COMTIs vs. placebo n=3 MAOBIs vs. placebo n=2	Not serious ¹	Not serious	Serious ²	Not serious	Moderate			
DAs vs. DAs n=2 COMTIs vs. COMTIs n=1 DA vs. COMTI n=1								
¹ Individual studies at risk of bias, but or ² Considered serious as population is n	verall risk of bias rated lo ot as defined in protocol	w due to consistency of effect	between studies at high a	nd low risk of bias				

Any adverse events

Quality assessment					
Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Quality
Any AEs					
51 DAs vs. placebo n=25 COMTIs vs. placebo n=14	Not serious ¹	Not serious	Serious ²	Not serious	Moderate
MAOBIs vs. placebo n=6 DAs vs. DAs n=4 DA vs. COMTI n=1 COMTI vs. COMTI n=1					
¹ Individual studies at risk of bias, but o ² Considered serious as population is n	verall risk of bias rated lo ot as defined in protocol	w due to consistency of effect	between studies at high a	nd low risk of bias	

Adverse event discontinuations

Quality assessment								
Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Quality			
AE discontinuations								
58	Not serious ¹	Not serious	Serious ²	Not serious	Moderate			
DAs vs. placebo n=29								
COMTIs vs. placebo n=17								
MAOBIs vs. placebo n=5								
DAs vs. DAs n=4								
DAs vs. COMTIs n=2								
COMTI vs. COMTI n=1								
¹ Individual studies at risk of bias, but o	verall risk of bias rated lo	w due to consistency of effect	between studies at high a	nd low risk of bias				

²Considered serious as population is not as defined in protocol

Pairwise meta-analyses

Dopamine agonists vs. Placebo

Off time

	Dopamine agonists			Placebo Mean Difference				Mean Difference			
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Random, 95% Cl	IV, Random, 95% Cl		
'Cochrane 2010	-1.54	3.68712544	1438	0	3.68712544	1163	34.0%	-1.54 [-1.82, -1.26]			
Nicholas 2014	-2.1212	2.1927	397	-1.5	3.1	105	20.8%	-0.62 [-1.25, 0.01]			
Nomoto 2014	-2.1	3.1	54	-0.7	2.8	56	10.3%	-1.40 [-2.51, -0.29]			
Pahwa 2007	-2.1	4.5368	201	-0.3	4.4109	190	14.1%	-1.80 [-2.69, -0.91]			
Poewe 2007	-1.7596	2.8651	401	0	2.8559	100	20.9%	-1.76 [-2.39, -1.13]	_		
Total (95% CI) Heterogeneity: Tau ² :	: П 11 [.] Сhi ^a	*= 8 55 df = 4 (2491 P = 0.0	7):12 = 5	3%	1614	100.0%	-1.42 [-1.83, -1.01]	→		
Test for overall effect	Z = 6.75 (P < 0.00001)	, = 0.0	.,,	0,0				-2 -1 0 1 2 Favours dopamine agonists Favours placebo		

UPDRS II

	Dopamine agonists				Placebo			Mean Difference	Mean Difference	
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Random, 95% Cl	IV, Random, 95% Cl	
'Cochrane 2010	-2.05	4.30785441	574	0	4.30785441	440	17.4%	-2.05 [-2.58, -1.52]	_ 	
Mizuno 2003	-3.6115	3.8804	206	-2.03	3.35	107	12.8%	-1.58 [-2.41, -0.75]		
Mizuno 2007	-2.7	4	121	-1	3.2	120	11.7%	-1.70 [-2.61, -0.79]		
Nicholas 2014	-1.9742	4.3466	392	-0.9	3.7	105	12.8%	-1.07 [-1.90, -0.25]		
Nomoto 2014	-3.8	3.6	82	-1.6	2.6	86	11.2%	-2.20 [-3.15, -1.25]		
Pahwa 2007	-3.5	5.4739	197	-0.9	5.358	184	9.7%	-2.60 [-3.69, -1.51]		
Poewe 2007	-4.3985	4.4504	405	-2	4.3	101	11.3%	-2.40 [-3.34, -1.45]		
PSG 2007	-3.35	4.36	109	-2.77	5.21	35	4.3%	-0.58 [-2.49, 1.33]		
Watts 2010	-1.5	3.8	82	-1.2	3.9	83	8.8%	-0.30 [-1.47, 0.87]		
Total (95% CI)			2168			1261	100.0%	-1.72 [-2.16, -1.27]	◆	
Heterogeneity: Tau ² =	= 0.22; Chi	² = 16.21, df = 8	8 (P = 0.	04); I²=	51%				-4 -2 0 2 4	1
restior overall ellect.	2=7.61(P < 0.00001)							Favours dopamine agonists Favours placebo	

UPDRS III

	Dopamine agonists				Placebo Mean Difference			Mean Difference	Mean Difference		
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Random, 95% Cl	IV, Random, 95% Cl		
'Cochrane 2010	-4.86	9.8842617	774	0	9.8842617	629	16.0%	-4.86 [-5.90, -3.82]			
Mizuno 2003	-10.8564	10.1089	206	-5.55	8.08	107	11.7%	-5.31 [-7.37, -3.24]			
Mizuno 2007	-9.5	8.7	121	-4.5	7.9	120	11.5%	-5.00 [-7.10, -2.90]			
Nicholas 2014	-4.3087	7.8591	393	-2.5	8.2	105	13.0%	-1.81 [-3.56, -0.06]			
Nomoto 2014	-10.1	9	86	-4.4	7.4	86	10.1%	-5.70 [-8.16, -3.24]			
Pahwa 2007	-6.5	12.6052	194	-1.7	12.3779	183	9.9%	-4.80 [-7.32, -2.28]	-		
Poewe 2007	-9.4941	8.3474	405	-4.3	9.3	101	12.0%	-5.19 [-7.18, -3.21]			
PSG 2007	-6.92	9.3	109	-3.77	10.66	35	5.9%	-3.15 [-7.09, 0.79]			
Watts 2010	-3.7	9.3	83	-3.5	7	81	9.9%	-0.20 [-2.72, 2.32]			
Total (95% CI)			2371			1447	100.0%	-4.09 [-5.25, -2.92]	◆		
Heterogeneity: Tau ² =	: 1.92; Chi ²:	= 22.93, df = 8	8 (P = 0.)	003); l² :	= 65%						
Test for overall effect:	Z = 6.88 (P	< 0.00001)							Favours dopamine agonists Favours placebo		

PDQ-39

	Dopan	opamine agonists			Placebo			Mean Difference	Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Random, 95% Cl	IV, Random, 95% Cl
Poewe 2007	-4.8985	9.6467	405	-1.3	9.4	101	52.3%	-3.60 [-5.66, -1.54]	
Watts 2010	-2.4	7.2	66	-2.4	7.3	59	47.7%	0.00 [-2.55, 2.55]	
Total (95% CI)			471			160	100.0%	-1.88 [-5.40, 1.64]	
Heterogeneity: Tau ² = Test for overall effect:	5.08; Chř Z = 1.05 (² = 4.64, d P = 0.30)	f=1 (P	= 0.03);	² = 7	'8%			Favours dopamine agonists Favours placebo

PDQUALIF

	Dopam	Dopamine agonists Placebo				Mean Difference	Mean Difi	ference			
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Fixed, 95% Cl	IV, Fixed,	95% CI	
PSG 2007	-5.41	11.75	109	-2.19	8.74	35	100.0%	-3.22 [-6.86, 0.42]		-	
Total (95% Cl)			109			35	100.0%	-3.22 [-6.86, 0.42]		-	
Test for overall effect: 3	piicable Z = 1.73 ((P = 0.08))						-10 -5 Ó Favours dopamine agonists	5 Favours placebo))

COMTIs vs. Placebo

Off time

		COMTIS			Placebo		Mean Difference Mean Difference				
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Fixed, 95% Cl	IV, Fixed, 95% Cl		
'Cochrane 2010	-0.83	2.41805517	1138	0	2.41805517	922	95.4%	-0.83 [-1.04, -0.62]			
Fenelon 2003	-0.8633	2.903	99	-0.53666	3.13167	63	4.6%	-0.33 [-1.29, 0.64]			
									•		
Total (95% CI)			1237			985	100.0%	-0.81 [-1.01, -0.60]	◆		
Heterogeneity: Chi ² =	1.00, df =	1 (P = 0.32); l ^a	= 0%					-			
Test for overall effect:	Z=7.71 (P ≺ 0.00001)							Favours COMTIS Favours placebo		

UPDRS II

		COMTIS			Placebo			Mean Difference	Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Fixed, 95% Cl	IV, Fixed, 95% Cl
'Cochrane 2010	-0.91	3.95430496	944	0	3.95430496	697	86.5%	-0.91 [-1.30, -0.52]	
Fenelon 2003	-1.6	6.5133	198	-0.35	4.4573	126	9.1%	-1.25 [-2.45, -0.05]	
Tolosa 2014	9.1	3.6	39	11.2	4.2	42	4.5%	-2.10 [-3.80, -0.40]	
Total (95% CI)			1181			865	100.0%	-0.99 [-1.35, -0.63]	•
Heterogeneity: Chi² = Test for overall effect	: 1.98, df : Z = 5.41	= 2 (P = 0.37) (P < 0.00001	; ² = 09)	6					-4 -2 0 2 4 Favours COMTIs Favours placebo

UPDRS III

		COMTIS			Placebo			Mean Difference	Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Fixed, 95% Cl	IV, Fixed, 95% Cl
'Cochrane 2010	-2.02	7.53757201	1135	0	7.53757201	922	94.5%	-2.02 [-2.67, -1.37]	
Tolosa 2014	13.9	6.4	39	17.5	6.1	42	5.5%	-3.60 [-6.33, -0.87]	
Total (95% CI)			1174			964	100.0%	-2.11 [-2.74, -1.47]	•
Heterogeneity: Chi² = Test for overall effect:	1.22, df Z = 6.48	= 1 (P = 0.27) } (P < 0.00001	; I² = 18)	%					-4 -2 0 2 4 Favours COMTIs Favours placebo

PDQ-39

	COMTIS Placebo					Mean Difference		Mean	Differenc	e			
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Fixed, 95% Cl		IV, Fix	ed, 95% (1	
Tolosa 2014	146.3	24.4	39	139.4	25.9	42	100.0%	6.90 [-4.05, 17.85]					
Total (95% Cl)			39			42	100.0%	6.90 [-4.05, 17.85]					-
Heterogeneity: Not ap Test for overall effect:	Z = 1.23	9 8 (P = 0	0.22)						-20	-10 Favours COMTI	0 s Favou	10 rs placebo	20

MAOBIs vs. Placebo

Off time

		MAOBIS			Placebo			Mean Difference		Mean Di	fference		
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Fixed, 95% Cl		IV, Fixed	I, 95% CI		
'Cochrane 2010	-0.93	2.39190904	465	0	2.39190904	423	58.9%	-0.93 [-1.24, -0.62]					
Zhang 2013	-1.748	1.5027	119	-0.691	1.5031	125	41.1%	-1.06 [-1.43, -0.68]					
Total (95% CI)			584			548	100.0%	-0.98 [-1.22, -0.74]		•			
Heterogeneity: Chi ² =	0.26, df=	= 1 (P = 0.61);	I ² = 0%						-2	-1	 0	1	2
Test for overall effect	Z=7.96	(P < 0.00001)								Favours MAOBIs	Favours pl	acebo	

UPDRS II

	MAOBIS			Pla	acebo			Mean Difference	Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Fixed, 95% Cl	IV, Fixed, 95% Cl
Zhang 2013	-3.51	3.15	119	-1.66	2.94	125	100.0%	-1.85 [-2.62, -1.08]	
Total (95% CI)			119			125	100.0 %	-1.85 [-2.62, -1.08]	· • · ·
Heterogeneity: Not ap Test for overall effect:	Z = 4.74	(P < (0.0000 [,]	1)					-4 -2 0 2 4 Favours MAOBIs Favours placebo

UPDRS III

	М	MAOBIS Placebo						Mean Difference	Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Fixed, 95% Cl	IV, Fixed, 95% Cl
'Cochrane 2010	-3.4	7.45	222	-0.5	7.38	218	29.5%	-2.90 [-4.29, -1.51]	_
Zhang 2013	-4.45	3.53	119	-2.41	3.61	125	70.5%	-2.04 [-2.94, -1.14]	
Total (95% Cl)			341		,	343	100.0%	-2.29 [-3.05, -1.54]	. ◆
Heterogeneity: Chi ² = Test for overall effect:	1.04, df Z = 5.97	= 1 (P '(P < (= 0.31)).00001); * = 49)	6				-4 -2 0 2 4 Favours MAOBIs Favours placebo

Ropinirole – Placebo

Dyskinesia

	Ropini	role	Place	bo		Risk Ratio		Risk Ratio		
Study or Subgroup	Events	Total	Events	Total	Weight	M-H, Random, 95% Cl		M-H, Rand	om, 95% Cl	
'Cochrane 2010	67	320	17	268	24.6%	3.30 [1.99, 5.48]				
Lieberman 1997	32	95	7	54	23.2%	2.60 [1.23, 5.48]				
Mizuno 2007	14	121	2	122	18.0%	7.06 [1.64, 30.39]				
Mizuno 2014	23	167	1	85	14.3%	11.71 [1.61, 85.21]				
Watts 2010	3	104	18	104	20.0%	0.17 [0.05, 0.55]				
Total (95% CI)		807		633	100.0%	2.36 [0.77, 7.22]		-		
Total events	139		45							
Heterogeneity: Tau² =	1.26; Ch	i² = 25.	36, df = 4	(P < 0.	0001); i ² :	= 84%		01		100
Test for overall effect:	Z=1.50	(P = 0.1	3)				0.01	Favours ropinirole	Favours placebo	100

Hallucinations

	Ropini	role	Place	bo		Risk Ratio	Risk Ratio
Study or Subgroup	Events	Total	Events	Total	Weight	M-H, Fixed, 95% Cl	M-H, Fixed, 95% Cl
'Cochrane 2010	12	202	2	191	43.7%	5.67 [1.29, 25.02]	_
Mizuno 2007	12	121	2	122	42.3%	6.05 [1.38, 26.46]	— — —
Mizuno 2014	6	167	0	85	14.0%	6.65 [0.38, 116.75]	
Total (95% CI)		490		398	100.0%	5.97 [2.23, 16.02]	
Total events	30		4				
Heterogeneity: Chi ² =	0.01, df=	2 (P =	0.99); i ² =	= 0%			
Test for overall effect:	Z = 3.55	(P = 0.0	004)				Favours ropinirole Favours placebo

Mortality



AE discontinuation

	Ropinirol	le	Placel	bo		Risk Ratio	Risk Ratio
Study or Subgroup	Events T	Fotal I	Events	Total	Weight	M-H, Fixed, 95% Cl	M-H, Fixed, 95% Cl
'Cochrane 2010	37	366	26	290	46.3%	1.13 [0.70, 1.82]	
Mizuno 2007	13	121	14	122	22.3%	0.94 [0.46, 1.91]	
Mizuno 2014	13	167	8	85	16.9%	0.83 [0.36, 1.92]	
Watts 2010	15	105	9	104	14.4%	1.65 [0.76, 3.60]	
Total (95% CI)		759		601	100.0%	1.11 [0.80, 1.53]	
Total events	78		57				
Heterogeneity: Chi ² =	1.69, df = 3	8 (P = 0).64); I ² =	:0%		-	
Test for overall effect:	Z=0.63 (P	9 = 0.53	3)				Favours ropinirole Favours placebo

Any AEs

	Ropini	role	Placebo			Risk Ratio	Risk Ratio
Study or Subgroup	Events	Total	Events	Total	Weight	M-H, Fixed, 95% Cl	M-H, Fixed, 95% Cl
'Cochrane 2010	185	271	140	236	29.8%	1.15 [1.01, 1.31]	
Mizuno 2007	102	121	80	122	15.9%	1.29 [1.11, 1.49]	
Mizuno 2014	130	167	59	85	15.6%	1.12 [0.95, 1.32]	
Pahwa 2007	129	202	106	191	21.7%	1.15 [0.98, 1.36]	
Watts 2010	91	104	86	104	17.1%	1.06 [0.94, 1.19]	
Total (95% CI)		865		738	100.0%	1.15 [1.08, 1.23]	-
Total events	637		471				
Heterogeneity: Chi ² =	4.29, df=	4 (P =	0.37); l² =	= 7%			0.7 0.85 1 1.2 1.5
Test for overall effect:	Z = 4.15	(P < 0.0	JOO1)				Favours ropinirole Favours placebo

SAEs

	Ropini	role	Place	bo		Risk Ratio	Risk Ratio	
Study or Subgroup	Events	Total	Events	Total	Weight	M-H, Fixed, 95% Cl	M-H, Fixed, 95% Cl	
Mizuno 2007	6	121	3	122	11.5%	2.02 [0.52, 7.88]		
Mizuno 2014	5	167	6	85	30.6%	0.42 [0.13, 1.35]		
Watts 2010	15	105	15	104	57.9%	0.99 [0.51, 1.92]		
Total (95% CI)		393		311	100.0%	0.94 [0.56, 1.57]	-	
Total events	26		24					
Heterogeneity: Chi² =	3.04, df=	2 (P =	0.22); l² :	= 34%				10
Test for overall effect:	Z=0.25	(P = 0.8	30)				Favours ropinirole Favours placebo	.0

Psychosis (PPRS)

	Ropinirole Placebo)		Mean Difference	Mean Difference				
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Fixed, 95% CI	IV, Fixed, 95% Cl		
Watts 2010	0.3	1.7	83	0	1.6	82	100.0%	0.30 [-0.20, 0.80]			
Total (95% CI)			83			82	100.0%	0.30 [-0.20, 0.80]			
Heterogeneity: Not ap Test for overall effect:	Z = 1.17	' (P =	0.24)						-2 -1 0 1 2 Favours ropinirole Favours placebo		

Rotigotine - Placebo

	Rotigo	tine	Place	bo		Risk Ratio	Risk Ratio
Study or Subgroup	Events	Total	Events	Total	Weight	M-H, Fixed, 95% Cl	M-H, Fixed, 95% Cl
LeWitt 2007	35	229	8	120	38.1%	2.29 [1.10, 4.78]	
Mizuno 2014	27	168	1	85	4.8%	13.66 [1.89, 98.82]	│ ——→
Nicholas 2014	34	406	3	108	17.2%	3.01 [0.94, 9.63]	
Nomoto 2014	12	87	7	87	25.4%	1.71 [0.71, 4.15]	- -
Poewe 2007	24	204	3	101	14.6%	3.96 [1.22, 12.84]	· · · · · · · · · · · · · · · · · · ·
Total (95% CI)		1094		501	100.0%	3.06 [1.95, 4.81]	-
Total events	132		22				
Heterogeneity: Chi ² =	4.63, df=	4 (P =	0.33); l² =	= 14%			
Test for overall effect:	Z= 4.85	(P < 0.0)0001)				Favours rotigotine Favours placebo

Hallucinations

	Rotigotine	e Place	bo		Risk Ratio	Risk Ratio
Study or Subgroup	Events To	otal Events	Total	Weight	M-H, Fixed, 95% Cl	M-H, Fixed, 95% Cl
LeWitt 2007	23 2	229 3	120	41.4%	4.02 [1.23, 13.11]	_
Mizuno 2014	3 ′	168 0	85	7.0%	3.56 [0.19, 68.18]	
Nicholas 2014	10 4	406 1	108	16.6%	2.66 [0.34, 20.55]	
Nomoto 2014	8	87 2	87	21.0%	4.00 [0.87, 18.30]	
Poewe 2007	10 2	204 1	101	14.1%	4.95 [0.64, 38.14]	
Total (95% CI)	10)94	501	100.0%	3.89 [1.82, 8.30]	-
Total events	54	7				
Heterogeneity: Chi ² =	0.19, df = 4	(P = 1.00); I ² :	= 0%			
Test for overall effect:	Z = 3.51 (P =	= 0.0004)				Favours rotigotine Favours placebo

AE discontinuation

	Rotigo	tine	Place	bo		Risk Ratio	Risk Ratio				
Study or Subgroup	Events	Total	Events	Total	Weight	M-H, Fixed, 95% Cl	M-H, Fixed, 95% Cl				
LeWitt 2007	18	231	11	120	21.6%	0.85 [0.42, 1.74]					
Mizuno 2014	13	168	8	85	15.8%	0.82 [0.35, 1.91]					
Nicholas 2014	44	408	17	108	40.1%	0.69 [0.41, 1.15]					
Nomoto 2014	12	87	7	87	10.4%	1.71 [0.71, 4.15]					
Poewe 2007	11	205	6	99	12.1%	0.89 [0.34, 2.32]					
Total (95% Cl)		1099		499	100.0%	0.87 [0.63, 1.21]	-				
Total events	98		49								
Heterogeneity: Chi ² =	3.11, df=	: 4 (P =	0.54); l² :	= 0%							
Test for overall effect:	Z = 0.82	(P = 0.4	11)				Favours rotigotine Favours placebo				

SAEs

	Rotigo	tine	Place	bo		Risk Ratio	Risk Ratio
Study or Subgroup	Events	Total	Events	Total	Weight	M-H, Fixed, 95% Cl	M-H, Fixed, 95% Cl
Mizuno 2014	7	168	6	85	42.2%	0.59 [0.20, 1.70]	_
Nicholas 2014	9	408	5	108	41.9%	0.48 [0.16, 1.39]	
Nomoto 2014	3	87	3	87	15.9%	1.00 [0.21, 4.82]	
Total (95% CI)		663		280	100.0%	0.61 [0.31, 1.19]	
Total events	19		14				
Heterogeneity: Chi ² =	0.59, df=	: 2 (P =	0.75); l² :	= 0%			
Test for overall effect:	Z=1.45	(P = 0.1	5)				Favours rotigotine Favours placebo

Any AEs

	Rotigot	tine	Place	bo		Risk Ratio	Risk Ratio
Study or Subgroup	Events	Total	Events	Total	Weight	M-H, Random, 95% Cl	M-H, Random, 95% Cl
Mizuno 2014	149	168	59	85	21.8%	1.28 [1.10, 1.49]	
Nicholas 2014	295	406	78	108	25.3%	1.01 [0.88, 1.15]	_
Nomoto 2014	82	87	77	87	33.9%	1.06 [0.97, 1.17]	
Poewe 2007	141	204	65	99	19.0%	1.05 [0.89, 1.25]	
Total (95% CI)		865		379	100.0%	1.09 [0.99, 1.20]	-
Total events	667		279				
Heterogeneity: Tau² =	0.00; Chi	i² = 6.0	7, df = 3 (P = 0.1	1); I² = 51	%	
Test for overall effect:	Z=1.80 ((P = 0.0)7)				Favours rotigotine Favours placebo

Mortality

	Rotigo	Rotigotine Placebo				Risk Ratio	Risk Ratio
Study or Subgroup	Events	Total	Events	Total	Weight	M-H, Fixed, 95% Cl	M-H, Fixed, 95% Cl
Nicholas 2014	2	406	0	108	100.0%	1.34 [0.06, 27.69]	
Total (95% CI)		406		108	100.0%	1.34 [0.06, 27.69]	
Total events	2		0				
Heterogeneity: Not ap Test for overall effect:	plicable Z = 0.19	(P = 0.8	35)				0.005 0.1 1 10 200 Eavours rotigotine Eavours placebo

ICD

	Rotigo	tine	Place	bo		Risk Ratio	Risk Ratio
Study or Subgroup	Events	Total	Events	Total	Weight	M-H, Fixed, 95% Cl	M-H, Fixed, 95% Cl
Nicholas 2014	5	408	0	108	100.0%	2.93 [0.16, 52.61]	
Total (95% CI)		408		108	100.0%	2.93 [0.16, 52.61]	
Total events	5		0				
Heterogeneity: Not ap Test for overall effect:	plicable Z = 0.73	(P = 0.4	17)				0.005 0.1 1 10 200 Favours rotigotine Favours placebo

Pramipexole vs. Placebo

	Pramipe	exole	Place	bo		Risk Ratio	Risk Ratio
Study or Subgroup	Events	Total	Events	Total	Weight	M-H, Fixed, 95% Cl	M-H, Fixed, 95% Cl
'Cochrane 2010	233	706	118	620	78.6%	1.73 [1.43, 2.10]	-
Mizuno 2003	16	102	6	108	3.6%	2.82 [1.15, 6.93]	
Poewe 2007	31	201	3	101	2.5%	5.19 [1.63, 16.58]	
PSG 2007	23	109	4	35	3.8%	1.85 [0.69, 4.97]	
Schapira 2011	59	339	14	178	11.5%	2.21 [1.27, 3.85]	
Total (95% Cl)		1457		1042	100.0%	1.92 [1.61, 2.29]	•
Total events	362		145				
Heterogeneity: Chi ² =	4.85, df =	4 (P = 0	.30); i² = 1	18%			
Test for overall effect:	Z = 7.29 (P < 0.00	001)				Favours pramipexole Favours placebo

Hallucinations

	Pramipe	exole	Place	bo		Risk Ratio	Risk Ratio
Study or Subgroup	Events	Total	Events	Total	Weight	M-H, Fixed, 95% Cl	M-H, Fixed, 95% Cl
'Cochrane 2010	84	658	30	502	82.4%	2.14 [1.43, 3.19]	
Mizuno 2003	14	102	4	108	9.4%	3.71 [1.26, 10.89]	
Poewe 2007	14	201	1	101	3.2%	7.03 [0.94, 52.75]	
PSG 2007	18	109	0	35	1.8%	12.11 [0.75, 195.91]	
Schapira 2011	18	339	1	178	3.2%	9.45 [1.27, 70.22]	
Total (95% CI)		1409		924	100.0%	2.86 [1.99, 4.09]	•
Total events	148		36				
Heterogeneity: Chi ² =	5.42, df =	4 (P = 0	.25); I² =	26%			
Test for overall effect:	Z = 5.72 (ł	● < 0.00	001)				Favours pramipexole Favours placebo

Mortality

	Pramipe	exole	Place	bo		Risk Ratio			Risk Ratio				
Study or Subgroup	Events	Total	Events	Total	Weight	M-H, Random, 95% Cl			M-H, Rand	om, 95%	CI		
'Cochrane 2010	0	120	0	131		Not estimable							
Total (95% CI)		120		131		Not estimable							
Total events	0		0										
Heterogeneity: Not ap Test for overall effect:	plicable Not applic	able					0.1	0.2 Favours	0.5 pramipexole	1 2 Favour	2 2 s placebo	5	10

AE discontinuations

	Pramipe	exole	Place	bo		Risk Ratio	Risk Ratio
Study or Subgroup	Events	Total	Events	Total	Weight	M-H, Fixed, 95% Cl	M-H, Fixed, 95% Cl
'Cochrane 2010	67	675	76	590	76.4%	0.77 [0.57, 1.05]	
Mizuno 2003	8	102	8	108	7.3%	1.06 [0.41, 2.72]	
Poewe 2007	14	202	6	99	7.6%	1.14 [0.45, 2.89]	
Schapira 2011	16	339	7	178	8.7%	1.20 [0.50, 2.86]	
Total (95% Cl)		1318		975	100.0%	0.86 [0.66, 1.12]	-
Total events	105		97				
Heterogeneity: Chi ² =	1.60, df=	3 (P = 0	.66); I ² = I	0%			
Test for overall effect: Z = 1.14 (P = 0.26)							Favours pramipexole Favours placebo

Any AEs

	Pramipe	xole	Place	bo		Risk Ratio	Risk Ratio
Study or Subgroup	Events	Total	Events	Total	Weight	M-H, Fixed, 95% Cl	M-H, Fixed, 95% Cl
'Cochrane 2010	370	452	280	367	50.9%	1.07 [1.00, 1.15]	
Mizuno 2003	87	102	83	108	13.3%	1.11 [0.97, 1.27]	
Poewe 2007	140	202	65	99	14.4%	1.06 [0.89, 1.25]	
Schapira 2011	202	339	99	178	21.4%	1.07 [0.91, 1.25]	
Total (95% Cl)		1095		752	100.0%	1.08 [1.01, 1.14]	◆
Total events	799		527				
Heterogeneity: Chi ² =	0.28, df=	3 (P = 0	.96); I² = I	0%			
Test for overall effect: $Z = 2.44$ (P = 0.01)							Favours pramipexole Favours placebo

SAEs

	Pramipe	xole	Place	bo		Risk Ratio	Risk Ratio
Study or Subgroup	Events	Total	Events	Total	Weight	M-H, Fixed, 95% Cl	M-H, Fixed, 95% Cl
Mizuno 2003	3	102	0	107	5.4%	7.34 [0.38, 140.36]	
PSG 2007	3	109	0	35	8.3%	2.29 [0.12, 43.30]	•
Schapira 2011	12	339	6	178	86.4%	1.05 [0.40, 2.75]	
Total (95% Cl)		550		320	100.0%	1.49 [0.64, 3.44]	
Total events	18		6				
Heterogeneity: Chi ² =	1.71, df=	2 (P = 0	.43); I² =				
Test for overall effect:	Z = 0.93 (I	P = 0.35	i)		Favours pramipexole Favours placebo		

Cabergoline vs. Placebo

	Caberge	Cabergoline Placebo			Risk Ratio	Risk Ratio	
Study or Subgroup	Events	Total	Events	Total	Weight	M-H, Fixed, 95% Cl	I M-H, Fixed, 95% Cl
'Cochrane 2010	135	291	57	158	100.0%	1.29 [1.01, 1.64]	
Total (95% CI)		291		158	100.0%	1.29 [1.01, 1.64]	
Total events	135		57				
Heterogeneity: Not ap	plicable	n - o o.	0				0.5 0.7 1 1.5 2
rest for overall effect.	Z = 2.04 (P = 0.04	¥)				Favours cabergoline Favours placebo

Hallucinations

	Cabergoline Placebo			Risk Ratio	Risk Ratio		
Study or Subgroup	Events	Total	Events	Total	Weight	M-H, Fixed, 95% Cl	M-H, Fixed, 95% Cl
'Cochrane 2010	14	165	4	103	100.0%	2.18 [0.74, 6.46]	
Total (95% Cl)		165		103	100.0%	2.18 [0.74, 6.46]	
Total events	14		4				
Heterogeneity: Not ap	plicable						
Test for overall effect: Z = 1.41 (P = 0.16)							Favours cabergoline Favours placebo

Mortality

	Cabergoline Placebo		Risk Ratio		Risk Ratio		
Study or Subgroup	Events	Total	Events	Total	Weight	M-H, Fixed, 95% Cl	M-H, Fixed, 95% Cl
'Cochrane 2010	0	20	1	20	100.0%	0.33 [0.01, 7.72]	
Total (95% CI)		20		20	100.0%	0.33 [0.01, 7.72]	
Total events	0		1				
Heterogeneity: Not ap Test for overall effect:	plicable Z = 0.69 (P = 0.49	3)				0.005 0.1 1 10 200 Favours cabergoline Favours placebo

AE discontinuation

	Cabergoline Placebo			Risk Ratio	Risk Ratio		
Study or Subgroup	Events	Total	Events	Total	Weight	M-H, Fixed, 95% Cl	M-H, Fixed, 95% Cl
'Cochrane 2010	12	165	6	103	100.0%	1.25 [0.48, 3.22]	
Total (95% CI)		165		103	100.0%	1.25 [0.48, 3.22]	
Total events	12		6				
Heterogeneity: Not ap	plicable						
Test for overall effect:	Test for overall effect: Z = 0.46 (P = 0.65)						Favours cabergoline Favours placebo

Any AEs

	Cabergoline Placebo			Risk Ratio	Risk Ratio			
Study or Subgroup	Events	Total	Events	Total	Weight	M-H, Fixed, 95% Cl	M-H, Fixed, 95% Cl	
'Cochrane 2010	141	165	75	103	100.0%	1.17 [1.03, 1.34]		
Total (95% CI)		165		103	100.0%	1.17 [1.03, 1.34]		
Total events	141		75					
Heterogeneity: Not ap Test for overall effect:	: applicable ect: Z = 2.35 (P = 0.02)						0.7 0.85 1 1.2 1.5 Favours cabergoline Favours placebo	2

Bromocriptine vs. Placebo

Dyskinesia

	Bromocri	iptine	ntine Placebo			Risk Ratio	Risk Ratio
Study or Subgroup	Events	Total	Events	Total	Weight	M-H, Fixed, 95% Cl	M-H, Fixed, 95% Cl
'Cochrane 2010	45	198	23	191	79.8%	1.89 [1.19, 2.99]	
Mizuno 2003	9	105	6	108	20.2%	1.54 [0.57, 4.18]	
Total (95% CI)		303		299	100.0%	1.82 [1.20, 2.76]	
Total events	54		29				
Heterogeneity: Chi ² =	0.13, df = 1	(P = 0.	72); I ^z = 0	%			
Test for overall effect:	Z = 2.80 (P	= 0.005	5)				Favours bromocriptine Favours placebo

Hallucinations

	Вготосгі	ptine	Placebo		Risk Ratio		Risk Ratio		
Study or Subgroup	Events	Total	Events	Total	Weight	M-H, Random, 95% Cl	M-H, Rand	om, 95% Cl	
'Cochrane 2010	17	198	16	191	54.6%	1.02 [0.53, 1.97]			
Mizuno 2003	16	105	4	108	45.4%	4.11 [1.42, 11.90]			_
Total (95% Cl)		303		299	100.0%	1.93 [0.49, 7.56]			
Total events	33		20						
Heterogeneity: Tau ² =	0.78; Chi ² :	= 4.85, (df = 1 (P =	= 0.03);	l² = 79%				20
Test for overall effect:	Z=0.94 (P	= 0.35)					Favours bromocriptine	Favours placebo	20

AE discontinuation

	Bromocr	Bromocriptine Placebo			Risk Ratio	Risk Ratio	
Study or Subgroup	Events	Total	Events	Total	Weight	M-H, Fixed, 95% Cl	M-H, Fixed, 95% Cl
'Cochrane 2010	39	232	40	220	83.9%	0.92 [0.62, 1.38]	—— — ——
Mizuno 2003	12	105	8	108	16.1%	1.54 [0.66, 3.62]	
Total (95% Cl)		337		328	100.0%	1.02 [0.71, 1.47]	-
Total events	51		48				
Heterogeneity: Chi ² =	1.14, df = 1	(P = 0.)	29); I ^z = 1	2%			
Test for overall effect:	Z=0.13 (P	9 = 0.90)					Favours bromocriptine Favours placebo

Any AEs

	Bromocriptine Placebo			Risk Ratio	Risk Ratio		
Study or Subgroup	Events	Total	Events	Total	Weight	M-H, Fixed, 95% Cl	M-H, Fixed, 95% Cl
'Cochrane 2010	54	114	44	108	35.6%	1.16 [0.86, 1.57]	
Mizuno 2003	95	105	83	108	64.4%	1.18 [1.04, 1.33]	
Total (95% CI) Total events	149	219	127	216	100.0 %	1.17 [1.03, 1.34]	-
Heterogeneity: Chi ² = Test for overall effect:	0.01, df = 1 Z = 2.37 (P	(P = 0.9 = 0.02)	93); I² = 0	0.7 0.85 1 1.2 1.5 Favours bromocriptine Favours placebo			

SAEs

	Bromocri	Bromocriptine		Bromocriptine Placebo		bo	Risk Ratio		Risk Ratio
Study or Subgroup	Events	vents Total Events Total			Weight	M-H, Fixed, 95% Cl	M-H, Fixed, 95% Cl		
Mizuno 2003	0	104	0	107		Not estimable			
Total (95% CI)		104		107		Not estimable			
Total events	0		0						
Heterogeneity: Not ap Test for overall effect:	plicable Not applica	able					0.1 0.2 0.5 1 2 5 10 Favours bromocriptine Favours placebo		

Pergolide vs. Placebo

Dyskinesia

	Pergol	ide	de Placebo			Risk Ratio	Risk Ratio		
Study or Subgroup	Events	Total	Events	Total	Weight	M-H, Fixed, 95% Cl	M-H, Fixed, 95% Cl		
'Cochrane 2010	118	189	46	187	100.0%	2.54 [1.93, 3.34]			
Total (95% CI)		189		187	100.0%	2.54 [1.93, 3.34]	•		
Total events	118		46						
Heterogeneity: Not ap	plicable								
Test for overall effect: Z = 6.66 (P < 0.00001)							Favours pergolide Favours placebo		

Hallucinations

	Pergolide Placebo		Risk Ratio			Risk Ratio						
Study or Subgroup	Events	Total	Events	Total	Weight	M-H, Fixed, 95% Cl			M-H, Fixe	d, 95% Cl		
'Cochrane 2010	26	189	6	187	100.0%	4.29 [1.81, 10.18]						
Total (95% CI)		189		187	100.0%	4.29 [1.81, 10.18]						
Total events	26		6									
Heterogeneity: Not ap Test for overall effect:	plicable Z = 3.30	(P = 0.0)010)				0.1	0.2 Favours	0.5 pergolide	l 1 1 2 Favours pla	5 cebo	10

Mortality

	Pergolide		Placebo			Risk Ratio	Risk Ratio	
Study or Subgroup	Events	Total	Events	Total	Weight	M-H, Fixed, 95% Cl	M-H, Fixed, 95% Cl	
'Cochrane 2010	1	189	2	187	100.0%	0.49 [0.05, 5.41]		
Total (95% CI)		189		187	100.0%	0.49 [0.05, 5.41]		
Total events	1		2					
Heterogeneity: Not ap Test for overall effect:	plicable Z = 0.58	(P = 0.6	i6)				0.05 0.2 1 5 20 Favours pergolide Favours placebo	

AE discontinuation

	Pergolide Placebo			Risk Ratio	Risk Ratio		
Study or Subgroup	Events	Total	Events	Total	Weight	M-H, Fixed, 95% Cl	M-H, Fixed, 95% Cl
'Cochrane 2010	18	189	8	187	100.0%	2.23 [0.99, 4.99]	
Total (95% CI)		189		187	100.0%	2.23 [0.99, 4.99]	
Total events	18		8				
Heterogeneity: Not ap	plicable						
Test for overall effect: Z = 1.94 (P = 0.05)							Favours pergolide Favours placebo

Entacapone vs. Placebo

Dyskinesia

	Entaca	Entacapone Placebo			Risk Ratio	Risk Ratio		
Study or Subgroup	Events	Total	Events	Total	Weight	M-H, Fixed, 95% Cl	M-H, Fixed, 95% Cl	
'Cochrane 2010	327	1430	122	1057	93.5%	1.98 [1.63, 2.40]		
Fenelon 2003	31	99	8	63	6.5%	2.47 [1.21, 5.02]		
Total (95% Cl)		1529		1120	100.0%	2.01 [1.67, 2.42]	•	
Total events	358		130					
Heterogeneity: Chi ² =	0.34, df=	1 (P = 0	0.56); I ^z =	0%				-
Test for overall effect:	Z=7.39 (P < 0.0	0001)				Favours entacapone Favours placebo	

Hallucinations

	Entaca	one	Place	bo		Risk Ratio	Risk Ratio
Study or Subgroup	Events	Total	Events	Total	Weight	M-H, Random, 95% Cl	M-H, Random, 95% Cl
'Cochrane 2010	36	924	23	740	62.5%	1.25 [0.75, 2.10]	
Fenelon 2003	0	99	4	63	37.5%	0.07 [0.00, 1.30]	<
Total (95% CI)		1023		803	100.0%	0.43 [0.03, 6.84]	
Total events	36		27				
Heterogeneity: Tau² =	3.14; Chi	z = 3.77	, df = 1 (F	P = 0.05	5); I² = 739	%	
Test for overall effect:	Z=0.60 (P = 0.5	5)				Favours entacapone Favours placebo

Mortality

	Entacap	Entacapone Placebo			Risk Ratio	Risk Ratio		
Study or Subgroup	Events	Total	Events	Total	Weight	M-H, Fixed, 95% Cl	M-H, Fixed, 95% Cl	
'Cochrane 2010	3	317	4	171	100.0%	0.40 [0.09, 1.79]		
Total (95% Cl)		317		171	100.0 %	0.40 [0.09, 1.79]		
Total events	3		4					
Heterogeneity: Not ap	plicable							100
Test for overall effect:	Z=1.19 (P = 0.23	3)				Favours entacapone Favours place	bo

AE discontinuation

	Entacap	one	Place	bo		Risk Ratio	Risk Ratio	
Study or Subgroup	Events	Total	Events	Total	Weight	M-H, Fixed, 95% Cl	M-H, Fixed, 95% Cl	
'Cochrane 2010	132	1232	71	917	88.9%	1.38 [1.05, 1.82]		
Destee 2009	12	110	0	66	0.7%	15.09 [0.91, 250.74]		٠
Fenelon 2003	17	99	7	63	9.3%	1.55 [0.68, 3.51]		
Tolosa 2014	3	45	1	49	1.0%	3.27 [0.35, 30.28]		٠
Total (95% Cl)		1486		1095	100.0%	1.51 [1.17, 1.95]	◆	
Total events	164		79					
Heterogeneity: Chi ² =	3.43, df=	3 (P = 0	0.33); I² =	13%				ł.
Test for overall effect:	Z = 3.16 (P = 0.0	02)				Favours entacapone Favours placebo	

Any AEs

	Entaca	pone	Place	bo		Risk Ratio	Risk Ratio
Study or Subgroup	Events	Total	Events	Total	Weight	M-H, Random, 95% Cl	M-H, Random, 95% Cl
'Cochrane 2010	716	1029	431	764	44.9%	1.23 [1.15, 1.33]	•
Destee 2009	32	110	5	68	7.3%	3.96 [1.62, 9.66]	_
Fenelon 2003	68	99	30	63	29.6%	1.44 [1.08, 1.93]	 − ∎ −
Tolosa 2014	20	45	19	49	18.2%	1.15 [0.71, 1.85]	
Total (95% Cl)		1283		944	100.0%	1.39 [1.07, 1.81]	◆
Total events	836		485				
Heterogeneity: Tau² =	0.04; Chi	r = 7.82	, df = 3 (F	° = 0.05	5); I² = 629	%	
Test for overall effect:	Z= 2.45 ((P = 0.0)	1)				Favours entacapone Favours placebo

SAEs

	Entaca	one	Place	bo		Risk Ratio	Risk Ratio
Study or Subgroup	Events	Total	Events	Total	Weight	M-H, Fixed, 95% Cl	M-H, Fixed, 95% Cl
Destee 2009	2	110	1	66	11.9%	1.20 [0.11, 12.98]	
Fenelon 2003	9	99	6	63	69.9%	0.95 [0.36, 2.55]	
Tolosa 2014	1	45	2	49	18.2%	0.54 [0.05, 5.80]	• •
Total (95% Cl)		254		178	100.0%	0.91 [0.39, 2.12]	
Total events	12		9				
Heterogeneity: Chi ² =	0.24, df=	2 (P = 0	0.89); I ^z =	0%			
Test for overall effect:	Z=0.22 (P = 0.8	2)				Favours entacapone Favours placebo

Tolcapone vs. Placebo

Dyskinesia

	Tolcapone P		Place	bo	Risk Ratio		Risk Ra	tio
Study or Subgroup	Events	Total	Events	Total	Weight	M-H, Fixed, 95% Cl	M-H, Fixed,	95% CI
'Cochrane 2010	125	288	49	291	100.0%	2.58 [1.93, 3.44]		
Total (95% CI)		288		291	100.0%	2.58 [1.93, 3.44]		-
Total events	125		49					
Heterogeneity: Not applicable								1.5 2
Test for overall effect: Z = 6.46 (P < 0.00001)							Favours tolcapone Fi	avours placebo

Hallucinations

	Tolcapone Placebo				Risk Ratio		Risk Ratio		
Study or Subgroup	Events	Total	Events	Total	Weight	M-H, Fixed, 95% Cl	M	-H, Fixed, 95% Cl	
'Cochrane 2010	20	155	11	213	100.0%	2.50 [1.23, 5.06]			
Total (95% CI)		155		213	100.0%	2.50 [1.23, 5.06]			
Total events	20		11						
Heterogeneity: Not applicable Test for overall effect: Z = 2.54 (P = 0.01)							0.2 0.5 Favours tolc:	1 2 apone Favours pla	5 cebo

AE discontinuation

	Tolcapone Placebo				Risk Ratio	Risk Ratio		
Study or Subgroup	Events	Total	Events	Total	Weight	M-H, Fixed, 95% Cl	M-H, Fixed, 95% Cl	
'Cochrane 2010	32	268	22	271	100.0%	1.47 [0.88, 2.46]		
Total (95% Cl)		268		271	100.0%	1.47 [0.88, 2.46]		
Total events	32		22					
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.47 (P = 0.14)							Favours tolcapone Favours placebo	

Any AEs

	Tolcapone Placebo				Risk Ratio	Risk Ratio			
Study or Subgroup	Events	Total	Events	Total	Weight	M-H, Fixed, 95% Cl	M-H, Fixed, 95	5% CI	
'Cochrane 2010	184	208	155	213	100.0%	1.22 [1.10, 1.34]			
Total (95% CI)		208		213	100.0%	1.22 [1.10, 1.34]			
Total events	184		155						
Heterogeneity: Not ap	plicable						0.7 0.95 1	12 15	
Test for overall effect: Z = 4.00 (P < 0.0001)							Favours tolcapone Fav	ours placebo	

Rasagiline vs. Placebo

	Rasagiline Placebo			Risk Ratio			Risk Ratio	
Study or Subgroup	Events	Total	Events	Total	Weight	M-H, Fixed, 95% Cl		M-H, Fixed, 95% Cl
'Cochrane 2010	12	231	9	229	86.0%	1.32 [0.57, 3.08]		
Zhang 2013	0	119	1	124	14.0%	0.35 [0.01, 8.44]	←	
Total (95% CI)		350		353	100.0 %	1.19 [0.53, 2.65]		
Total events	12		10					
Heterogeneity: Chi ² =	0.63, df=	: 1 (P =	0.43); I² =	= 0%				
Test for overall effect:	(P = 0.8	i8)				0.1	Favours rasagiline Favours placebo	

Hallucinations

	Rasagiline Placebo			Risk Ratio			Risk Ratio						
Study or Subgroup	Events	Total	Events	Total	Weight	M-H, Fixed, 95% Cl			M-H, Fixe	d, 95% (CI		
'Cochrane 2010	5	231	3	229	100.0%	1.65 [0.40, 6.83]							
Total (95% CI)		231		229	100.0%	1.65 [0.40, 6.83]						-	
Total events	5		3										
Heterogeneity: Not applicable Test for overall effect: Z = 0.69 (P = 0.49)							⊢ 0.1	0.2 Eavour:	0.5 s rasadiline	Favour	s placebo	5	10

AE discontinuation

	Rasagiline Placebo					Risk Ratio	Risk Ratio
Study or Subgroup	Events	Total	Events	Total	Weight	M-H, Fixed, 95% Cl	M-H, Fixed, 95% Cl
'Cochrane 2010	7	231	11	229	65.4%	0.63 [0.25, 1.60]	
Zhang 2013	3	119	6	125	34.6%	0.53 [0.13, 2.05]	
Total (95% Cl)		350		354	100.0%	0.59 [0.28, 1.28]	
Total events	10		17				
Heterogeneity: Chi ² =	0.05, df=	1 (P =					
Test for overall effect: Z = 1.33 (P = 0.18)							Favours rasagiline Favours placebo

Any AEs

	Rasagi	asagiline Placebo				Risk Ratio	Risk Ratio		
Study or Subgroup	Events	Total	Events	Total	Weight	M-H, Fixed, 95% Cl	M-H, Fixed, 95% Cl		
'Cochrane 2010	193	380	191	388	90.2%	1.03 [0.90, 1.19]			
Zhang 2013	27	119	21	125	9.8%	1.35 [0.81, 2.25]			
T-4-1 (0/21) (0)		400		540	400.0%	4 00 10 00 4 001			
Total (95% CI)		499		513	100.0%	1.06 [0.93, 1.22]			
Total events	220		212						
Heterogeneity: Chi ² =	1.01, df=	1 (P =	-						
Test for overall effect:	Z=0.87 ((P = 0.3)	9)				Favours rasagiline Favours placebo		

SAEs

	Rasagiline Placebo				Risk Ratio	Risk Ratio		
Study or Subgroup	Events	Total	Events	Total	Weight	M-H, Fixed, 95% Cl	M-H, Fixed, 95% Cl	
Zhang 2013	1	119	1	125	100.0%	1.05 [0.07, 16.60]		
Total (95% CI)		119		125	100.0%	1.05 [0.07, 16.60]		
Total events	1		1					
Heterogeneity: Not applicable								100
Test for overall effect: Z = 0.03 (P = 0.97)							Favours rasagiline Favours placebo	100

Selegiline vs. Placebo

Dyskinesia

	Selegiline Placebo					Risk Ratio		Risk Ratio
Study or Subgroup	Events	Total	Events	Total	Weight	M-H, Fixed, 95% Cl		M-H, Fixed, 95% Cl
Ondo 2007	8	100	2	50	16.4%	2.00 [0.44, 9.07]		
'Cochrane 2010	9	70	13	64	83.6%	0.63 [0.29, 1.38]		
Total (95% Cl)		170		114	100.0%	0.86 [0.44, 1.69]		
Total events	17		15					
Heterogeneity: Chi ² =	1.79, df=	: 1 (P =	0.18); I² :	= 44%				
Test for overall effect: Z = 0.45 (P = 0.66)							0.1	Favours selegiline Favours placebo

Hallucinations

	Selegiline Placebo					Risk Ratio	Risk Ratio
Study or Subgroup	Events	Total	Events	Total	Weight	M-H, Fixed, 95% Cl	M-H, Fixed, 95% Cl
'Cochrane 2010	3	50	1	46	100.0%	2.76 [0.30, 25.60]	
Total (95% CI)		50		46	100.0%	2.76 [0.30, 25.60]	
Total events	3		1				
Heterogeneity: Not ap Test for overall effect:	plicable Z = 0.89 ((P = 0.3	37)				0.02 0.1 1 10 50 Favours selegiline Favours placebo

AE discontinuation

	Selegiline Placebo			bo		Risk Ratio	Risk Ratio			
Study or Subgroup	Events	Total	Events	Total	Weight	M-H, Random, 95% Cl	M-H, Random, 95% Cl			
'Cochrane 2010	3	144	3	95	60.7%	0.66 [0.14, 3.20]	∎ +			
Ondo 2007	7	100	0	50	39.3%	7.57 [0.44, 130.01]				
Total (95% CI)		244		145	100.0%	1.72 [0.14, 20.91]				
Total events	10		3							
Heterogeneity: Tau ² =	2.02; Ch	i ² = 2.4 [°]	7, df = 1 (P = 0.1	2); I² = 60	%	0.01 0.1 1 10	100		
rest for overall effect.	∠=0.43	(== 0.8	11)				Favours selegiline Favours placebo			

Any AEs

	Selegiline Placebo			Risk Ratio			Risk Ratio			
Study or Subgroup	Events	Total	Events	Total	Weight	M-H, Fixed, 95% Cl		M-H, Fixed, 95%	CI	
'Cochrane 2010	45	114	21	64	35.9%	1.20 [0.79, 1.83]				_
Ondo 2007	73	100	36	50	64.1%	1.01 [0.82, 1.25]			_	
Total (95% Cl)		214		114	100.0%	1.08 [0.88, 1.33]		-		
Total events	118		57							
Heterogeneity: Chi ² =	0.61, df=	: 1 (P =	0.43); l² =			1.5				
Test for overall effect: Z = 0.75 (P = 0.45)							0.0	Favours selegiline Favou	rs placebo	2

SAEs

	Selegiline Placebo				Risk Ratio	Risk Ratio	
Study or Subgroup	Events	Total	Events	Total	Weight	M-H, Fixed, 95% Cl	M-H, Fixed, 95% Cl
Ondo 2007	8	100	1	50	100.0%	4.00 [0.51, 31.10]	
Total (95% CI)		100		50	100.0%	4.00 [0.51, 31.10]	
Total events	8		1				
Heterogeneity: Not ap Test for overall effect:	plicable Z = 1.32	(P = 0.1	9)				0.01 0.1 1 10 100 Favours selegiline Favours placebo

Amantadine vs. Placebo

Hyperkinesia (CDRS)

	Amantadine Placebo				Mean Difference	Mean Difference			
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Fixed, 95% Cl	IV, Fixed, 95% Cl
da Silva-Junior 2005	6.8	4.9	9	13	11.5	9	100.0%	-6.20 [-14.37, 1.97]	
Total (95% Cl)			9			9	100.0%	-6.20 [-14.37, 1.97]	
Heterogeneity: Not app Test for overall effect: 2	ilicable (= 1.49 (P = 0.	14)						-20 -10 0 10 20 Favours amantadine Favours placebo

Dystonia (CDRS)

	Amantadine Placebo			0		Mean Difference	Mean Difference			
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Fixed, 95% Cl	IV, Fixed, 95% Cl	
da Silva-Junior 2005	4.1	3.6	9	4.5	4.3	9	100.0%	-0.40 [-4.06, 3.26]		
Total (95% CI)			9			9	100.0 %	-0.40 [-4.06, 3.26]		
Heterogeneity: Not app Test for overall effect: Z	iicaple := 0.21 (P = 0.	83)						-10 -5 0 5 10 Favours amantadine Favours placebo	

UPDRS II

	Amantadine Placet			acebo	0		Mean Difference	Mean Difference	
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Fixed, 95% Cl	IV, Fixed, 95% Cl
da Silva-Junior 2005	12.1	8	9	13.8	7.9	9	100.0%	-1.70 [-9.05, 5.65]	
Total (95% CI)			9			9	100.0%	-1.70 [-9.05, 5.65]	
Heterogeneity: Not app Test for overall effect: Z	licable (= 0.45 (P = 0.	65)						-20 -10 0 10 20 Favours amantadine Favours placebo

UPDRS III



Ropinirole vs. Rotigotine

Any AEs

	Ropinirole Rotigotine				Risk Ratio	Risk Ratio	
Study or Subgroup	Events	Total	Events	Total	Weight	M-H, Fixed, 95% Cl	M-H, Fixed, 95% Cl
Mizuno 2014	130	167	149	168	100.0%	0.88 [0.80, 0.97]	
Total (95% CI)		167		168	100.0%	0.88 [0.80, 0.97]	◆
Total events	130		149				
Heterogeneity: Not ap	plicable						
Test for overall effect:	Z = 2.63 ((P = 0.0)09)				Favours ropinirole Favours rotigotine

SAEs

	Ropinirole Rotigotine				Risk Ratio	Risk Ratio					
Study or Subgroup	Events	Total	Events	Total	Weight	M-H, Fixed, 95% Cl	M-H, Fixed, 95% Cl				
Mizuno 2014	5	167	7	168	100.0%	0.72 [0.23, 2.22]					
Total (95% CI)		167		168	100.0%	0.72 [0.23, 2.22]					
Total events	5		7								
Heterogeneity: Not ap Test for overall effect:	plicable Z = 0.57	(P = 0.5	57)				0.1 0.2 0.5 1 2 5 10 Favours ropinirole Favours rotigotine				

AE discontinuation

	Ropinirole Rotigotine				Risk Ratio	Risk Ratio							
Study or Subgroup	Events	Total	Events	Total	Weight	M-H, Fixed, 95% Cl			M-H, Fixe	d, 95% C	1		
Mizuno 2014	13	167	13	168	100.0%	1.01 [0.48, 2.10]							
Total (95% Cl)		167		168	100.0%	1.01 [0.48, 2.10]							
Total events	13		13										
Heterogeneity: Not ap Test for overall effect:	plicable Z = 0.02 ((P = 0.9	9)				⊢ 0.1	0.2 Favour	0.5 s ropinirole	Favours	5 s rotigotine	!	10

Hallucinations

	Ropinirole Rotigotine				Risk Ratio	Risk Ratio	
Study or Subgroup	Events Total Events Total			Weight	M-H, Fixed, 95% Cl	M-H, Fixed, 95% Cl	
Mizuno 2014	6	167	3	168	100.0%	2.01 [0.51, 7.91]	
Total (95% CI)		167		168	100.0%	2.01 [0.51, 7.91]	
Total events	6		3				
Heterogeneity: Not ap	plicable						
Test for overall effect:	Z = 1.00 ((P = 0.3	32)				Favours ropinirole Favours rotigotine

	Ropinirole Rotigotine				Risk Ratio	Risk Ratio							
Study or Subgroup	Events	Total	Events	Total	Weight	M-H, Fixed, 95% Cl			M-H, Fixe	ed, 95% (CI		
Mizuno 2014	23	167	27	168	100.0%	0.86 [0.51, 1.43]				—			
Total (95% CI)		167		168	100.0%	0.86 [0.51, 1.43]							
Total events	23		27										
Heterogeneity: Not ap Test for overall effect:	plicable Z = 0.59	(P = 0.5	56)				L.1	0.2 Favou	0.5 rs ropinirole	1 2 Favour	2 : s rotigotine	5 8	10

Ropinirole vs. Bromocriptine

Hallucinations



Dyskinesia

	Ropinirole Bromocriptine				Risk Ratio	Risk Ratio					
Study or Subgroup	Events	Total	Events	Total	Weight	M-H, Fixed, 95% Cl		M-H, Fixe	ed, 95% Cl		
Clarke 2001 (B)	14	169	10	174	100.0%	1.44 [0.66, 3.16]					
Total (95% CI)		169		174	100.0%	1.44 [0.66, 3.16]					
Total events	14		10								
Heterogeneity: Not ap Test for overall effect:	plicable Z = 0.91	(P = 0.3	36)				0.1 0.1	2 0.5 Favours ropinirole	1 1 Favours	bromocrip	tine

Pramipexole vs. Bromocriptine

	Pramipexole		Bromocriptine			Risk Ratio	Risk Ratio			
Study or Subgroup	Events	Total	Events	Total	Weight	M-H, Fixed, 95% Cl		M-H, Fixed, 95% Cl		
'Cochrane 2010	7	114	1	108	10.4%	6.63 [0.83, 53.01]				
Mizuno 2003	16	102	9	105	89.6%	1.83 [0.85, 3.95]				
Total (95% Cl)		216		213	100.0%	2.33 [1.14, 4.74]				
Total events	23		10							
Heterogeneity: Chi ² =	1.35, df = 1	1 (P = 0	.25); I² = 2	6%						
Test for overall effect:	Z = 2.33 (F	P = 0.02	:)				0.1	Favours pramipexole Favours bromocriptine		

SAEs

	Pramipexole Bro		Bromoci	iptine		Risk Ratio	Risk Ratio			
Study or Subgroup	Events	Total	Events	Total	Weight	M-H, Fixed, 95% Cl	M-H, Fixed, 95% Cl			
Mizuno 2003	3	102	0	104	100.0%	7.14 [0.37, 136.43]				
Total (95% CI)		102		104	100.0%	7.14 [0.37, 136.43]				
Total events	3		0							
Heterogeneity: Not ap Test for overall effect:	oplicable Z=1.31 (F	P = 0.19)				0.005 0.1 1 10 20 Favours pramipexole Favours bromocriptine	0		

AE discontinuation

	Pramipe	xole	Вготосг	iptine		Risk Ratio	Risk Ratio
Study or Subgroup	Events	Total	Events	Total	Weight	M-H, Fixed, 95% Cl	CI M-H, Fixed, 95% CI
Mizuno 2003	8	102	12	105	100.0%	0.69 [0.29, 1.61]	
Total (95% CI)		102		105	100.0%	0.69 [0.29, 1.61]	
Total events	8		12				
Heterogeneity: Not ap	oplicable						
Test for overall effect:	Z=0.87 (F	P = 0.39)				Favours pramipexole Favours bromocriptine

Any AEs

	Pramipe	xole	Вготосг	iptine		Risk Ratio	Risk Ratio
Study or Subgroup	Events	Total	Events	Total	Weight	M-H, Fixed, 95% Cl	M-H, Fixed, 95% Cl
Mizuno 2003	87	102	95	105	100.0%	0.94 [0.85, 1.04]	
Total (95% CI)		102		105	100.0%	0.94 [0.85, 1.04]	◆
Total events	87		95				
Heterogeneity: Not ap	oplicable						
Test for overall effect:	Z = 1.14 (F	P = 0.26)				Favours pramipexole Favours bromocriptine

Hallucinations

	Pramipe	xole	Bromocr	iptine		Risk Ratio	Risk Ratio
Study or Subgroup	Events	Total	Events	Total	Weight	M-H, Fixed, 95% Cl	M-H, Fixed, 95% Cl
Mizuno 2003	14	102	16	105	100.0%	0.90 [0.46, 1.75]	
Total (95% CI)		102		105	100.0%	0.90 [0.46, 1.75]	
Total events	14		16				
Heterogeneity: Not ap Test for overall effect:	oplicable Z = 0.31 (F	^D = 0.76	i)				0.1 0.2 0.5 1 2 5 10 Favours pramipexole Favours bromocriptine

Rotigotine vs. Pramipexole

Any AEs

	Rotigo	tine	Pramipe	exole		Risk Ratio		Rist	<pre>Ratio</pre>		
Study or Subgroup	Events	Total	Events	Total	Weight	M-H, Fixed, 95% Cl		M-H, Fix	ed, 95% Cl		
Poewe 2007	141	204	140	202	100.0%	1.00 [0.88, 1.14]					
Total (95% CI)		204		202	100.0%	1.00 [0.88, 1.14]					
Total events	141		140								
Heterogeneity: Not ap Test for overall effect:	plicable Z = 0.04 ((P = 0.9	97)				0.5	0.7 Favours rotigotine	1 9 Favours p	1.5 ramipexol	2 e

	Rotigo	tine	Pramipe	exole		Risk Ratio			Risk	Ratio			
Study or Subgroup	Events	Total	Events	Total	Weight	M-H, Fixed, 95% Cl			M-H, Fixe	d, 95% C	3		
Poewe 2007	24	204	31	201	100.0%	0.76 [0.46, 1.25]							
Total (95% CI)		204		201	100.0%	0.76 [0.46, 1.25]				-			
Total events	24		31										
Heterogeneity: Not ap Test for overall effect:	plicable Z = 1.07	(P = 0.2	28)				L.1	0.2 Favour	0.5 s rotigotine	i : Favour:	l 2 s pramipe:	1 5 xole	10

Hallucinations



AE discontinuation

	Rotigo	tine	Pramipe	exole		Risk Ratio			Risk Ratio		
Study or Subgroup	Events	Total	Events	Total	Weight	M-H, Fixed, 95% Cl		M-H	, Fixed, 95%	CI	
Poewe 2007	11	205	14	202	100.0%	0.77 [0.36, 1.66]					
Total (95% CI)		205		202	100.0%	0.77 [0.36, 1.66]					
Total events	11		14								
Heterogeneity: Not ap Test for overall effect:	plicable Z = 0.66	(P = 0 f	51)				⊢ 0.1	0.2 0.5	1	2 5	10
restion overall effect.	2 - 0.00	(i = 0.c	~~~					Favours rotid	otine Favou	rs pramipexole	Э

Pramipexole vs. Pergolide

Any AEs

	Pramipe	exole	Pergol	lide		Risk Ratio			Risk	Ratio			
Study or Subgroup	Events	Total	Events	Total	Weight	M-H, Fixed, 95% Cl			M-H, Fixe	ed, 95% C	I		
Rektorova 2003	13	22	14	19	100.0%	0.80 [0.52, 1.24]				-			
Total (95% CI)		22		19	100.0%	0.80 [0.52, 1.24]				-			
Total events	13		14										
Heterogeneity: Not ap Test for overall effect:	plicable Z = 0.98 (I	P = 0.32	!)				⊢ 0.1	0.2 Favours pra	0.5 mipexole	1 2 Favours	pergolide	5	10

AE discontinuation



Cabergoline vs. Bromocriptine

Hallucinations

	Caberge	oline	Bromocr	iptine		Risk Ratio	Risk Ratio					
Study or Subgroup	Events	Total	Events	Total	Weight	M-H, Fixed, 95% Cl	M-H, Fixed, 95% Cl					
Clarke 2001 (A)	53	520	41	528	100.0%	1.31 [0.89, 1.94]						
Total (95% CI)		520		528	100.0%	1.31 [0.89, 1.94]						
Total events	53		41									
Heterogeneity: Not ap Test for overall effect:	plicable Z = 1.37 (P = 0.17	7)				0.5 0.7 1 1.5 2 Favours cabergoline Favours bromocriptine					

	Cabergoline Bromocriptine					Risk Ratio	Risk Ratio					
Study or Subgroup	Events	Total	Events	Total	Weight	M-H, Fixed, 95% Cl	M-H, Fixed, 95% Cl					
Clarke 2001 (A)	66	520	45	528	100.0%	1.49 [1.04, 2.13]						
Total (95% CI)		520		528	100.0%	1.49 [1.04, 2.13]						
Total events	66		45									
Heterogeneity: Not ap	plicable											
Test for overall effect:	Z = 2.17 (P = 0.00	3)				U.5 U.7 1 1.5 2 Favours cabergoline Favours bromocriptine					

Dopamine Agonists vs. COMTIs

UPDRS II

	Dopamine agonists							Mean Difference	Mean Difference			
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Fixed, 95% Cl	IV, Fixed, 95% Cl			
Deane 2004	-0.1	3.4	74	-0.9	4.2	72	49.8%	0.80 [-0.44, 2.04]				
Deuschl 2007	-2.5	3.9	69	-2.5	3.5	69	50.2%	0.00 [-1.24, 1.24]				
Total (95% Cl)			143			141	100.0%	0.40 [-0.48, 1.27]				
Heterogeneity: Chi ² = Test for overall effect:	0.80, df = 1 Z = 0.89 (P	(P = 0. = 0.37)	37); I² =)	0%					-2 -1 0 1 2 Favours dopamine agonists Favours COMTIs			

UPDRS III

	Dopamir	ne agor	nists	CC	MTIs	;		Mean Difference	Mean Difference				
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Fixed, 95% Cl		IV, Fix	ked, 95% Cl		
Deane 2004	-3.3	8.6	74	-3.1	8.5	72	50.0%	-0.20 [-2.97, 2.57]					
Deuschl 2007	-6.3	7.9	69	-6.3	8.7	69	50.0%	0.00 [-2.77, 2.77]			+		
Total (95% CI)			143			141	100.0%	-0.10 [-2.06, 1.86]					
Heterogeneity: Chi ² =	0.01, df = 1	(P = 0	.92); l² =	0%					-4	-2	0	2	4
lest for overall effect:	Z = 0.10 (P	= 0.92)						Favours d	opamine agonis	ts Favours C	OMTIs:	

Off time

	Dopamine agonists COMTIs							Mean Difference	Mean Difference			
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Fixed, 95% Cl	IV, Fixed, 95% Cl			
Deane 2004	2.4	3.9	74	3	4.1	72	30.5%	-0.60 [-1.90, 0.70]				
Deuschl 2007	-1.7	2.4	64	-1.8	2.7	71	69.5%	0.10 [-0.76, 0.96]				
Total (95% CI)			138			143	100.0%	-0.11 [-0.83, 0.60]				
Heterogeneity: Chi ² =	0.78, df = 1	(P = 0.	.38); I² =	0%								
Test for overall effect:	Z = 0.31 (P	= 0.76)						Favours dopaine agonists Favours COMTIs			

PDQ-39



Cabergoline vs. Entacapone

Any AEs

	Cabergoline		Entacapone		Risk Ratio			Risk Ratio			
Study or Subgroup	Events	Total	Events	Total	Weight	M-H, Fixed, 95% Cl		M-H, Fix	ed, 95% Cl		
Deuschl 2007	42	79	44	82	100.0%	0.99 [0.74, 1.32]					
Total (95% CI)		79		82	100.0%	0.99 [0.74, 1.32]					
Total events	42		44								
Heterogeneity: Not applicable Test for overall effect: Z = 0.06 (P = 0.95)							L.5	0.7 Favours cabergoline	1 Favours entac:	1.5 apone	2

SAEs

	Cabergoline		Entacapone		Risk Ratio			Risk Ratio		
Study or Subgroup	Events	Total	Events	Total	Weight	M-H, Fixed, 95% Cl	M-H, Fixed, 95% Cl			
Deuschl 2007	3	79	6	82	100.0%	0.52 [0.13, 2.00]				
Total (95% CI)		79		82	100.0%	0.52 [0.13, 2.00]				
Total events	3		6							
Heterogeneity: Not applicable Test for overall effect: Z = 0.95 (P = 0.34)							0.01	0.1 Favours cabergoline	1 10 Favours entacapone	100
AE discontinuation

	Caberge	oline	Entacap	one		Risk Ratio	Risk Ratio
Study or Subgroup	Events	Total	Events	Total	Weight	M-H, Fixed, 95% Cl	I M-H, Fixed, 95% Cl
Deuschl 2007	11	79	7	82	100.0%	1.63 [0.67, 4.00]	
Total (95% CI)		79		82	100.0%	1.63 [0.67, 4.00]	
Total events	11		7				
Heterogeneity: Not ap Test for overall effect:	plicable Z = 1.07 (P = 0.28	3)				0.1 0.2 0.5 1 2 5 10 Favours cabergoline Favours entacapone

Hallucinations

	Caberge	oline	Entacap	one		Risk Ratio		Risk	Ratio	
Study or Subgroup	Events	Total	Events	Total	Weight	M-H, Fixed, 95% Cl		M-H, Fixe	ed, 95% Cl	
Deuschl 2007	3	79	3	82	100.0%	1.04 [0.22, 4.99]				
Total (95% CI)		79		82	100.0%	1.04 [0.22, 4.99]				
Total events	3		3							
Heterogeneity: Not ap Test for overall effect:	plicable Z = 0.05 (P = 0.98	ō)				0.05	0.2 Favours cabergoline	1 5 Favours entaca;	20 Done

Bromocriptine vs. Tolcapone

Dyskinesia

	Bromocr	iptine	Tolcap	one		Risk Ratio	Risk Ratio
Study or Subgroup	Events	Total	Events	Total	Weight	M-H, Fixed, 95% Cl	M-H, Fixed, 95% Cl
Deane 2004	28	74	37	72	100.0%	0.74 [0.51, 1.06]	
Total (95% CI)		74		72	100.0%	0.74 [0.51, 1.06]	
Total events	28		37				
Heterogeneity: Not ap	oplicable						
Test for overall effect:	Z = 1.63 (F	'= 0.10)					Favours bromocriptine Favours tolcapone

Hallucinations



Pergolide vs. Tolcapone

Dyskinesia

	Pergol	lide	Tolcap	one		Risk Ratio		Ris	< Ratio	
Study or Subgroup	Events	Total	Events	Total	Weight	M-H, Fixed, 95% Cl		M-H, Fi>	ed, 95% Cl	
Deane 2004	25	102	34	71	100.0%	0.51 [0.34, 0.78]				
Total (95% CI)		102		71	100.0%	0.51 [0.34, 0.78]				
Total events	25		34							
Heterogeneity: Not ap Test for overall effect:	plicable Z = 3.14	(P = 0.0)02)				0.2	0.5 Favours pergolide	1 2 Favours tolcapone	5

AE discontinuation

	Pergo	lide	Tolcap	one		Risk Ratio	Risk Ratio
Study or Subgroup	Events	Total	Events	Total	Weight	M-H, Fixed, 95% Cl	M-H, Fixed, 95% Cl
Deane 2004	15	102	5	101	100.0%	2.97 [1.12, 7.87]	
Total (95% CI)		102		101	100.0%	2.97 [1.12, 7.87]	
Total events	15		5				
Heterogeneity: Not ap Test for overall effect:	plicable Z = 2.19	(P = 0.0)3)				0.1 0.2 0.5 1 2 5 10 Favours pergolide Favours tolcapone

Entacapone vs. Tolcapone

Any AEs



SAEs

	Entaca	one	Tolcap	one		Risk Ratio	Risk Ratio		
Study or Subgroup	Events	Total	Events	Total	Weight	M-H, Fixed, 95% Cl	M-H, Fixed, 95% Cl		
ESS 2007	1	75	6	75	100.0%	0.17 [0.02, 1.35]			
Total (95% CI)		75		75	100.0%	0.17 [0.02, 1.35]			
Total events	1		6						
Heterogeneity: Not ap Test for overall effect:	plicable Z = 1.68 (P = 0.0	9)				0.01 0.1 1 10 100 Favours entacapone Favours tolcapone		

AE discontinuation

	Entaca	one	Tolcap	one		Risk Ratio		Risk F	Ratio	
Study or Subgroup	Events	Total	Events	Total	Weight	M-H, Fixed, 95% Cl		M-H, Fixed	d, 95% Cl	
ESS 2007	1	75	0	75	100.0%	3.00 [0.12, 72.49]				-
Total (95% CI)		75		75	100.0%	3.00 [0.12, 72.49]				
Total events	1		0							
Heterogeneity: Not ap Test for overall effect:	plicable Z = 0.68 (P = 0.5	D)				0.005 0 Favours (I.1 1 entacapone	10 Favours tolcapone	200

Dyskinesia



Hallucinations

	Entacapone Tolcapone				Risk Ratio	Risk Ratio	
Study or Subgroup	Events	Total	Events	Total	Weight	M-H, Fixed, 95% Cl	M-H, Fixed, 95% Cl
ESS 2007	3	75	0	75	100.0%	7.00 [0.37, 133.22]	
Total (95% CI)		75		75	100.0%	7.00 [0.37, 133.22]	
Total events	3		0				
Heterogeneity: Not ap	plicable						
Test for overall effect:	Z=1.29 (P = 0.20	D)				Favours entacapone Favours tolcapone

Network meta-analyses

Efficacy outcomes by drug classes

Off time (hours) – FE model



Quantifying heterogeneity/inconsistency:

tau² = 0.0914; l² = 47.7%

Test of heterogeneity/inconsistency:

Q d.f. p.value

9.55 5 0.089

Differences between treatments – mean and 95% confidence interval

	Treatment A				
Treatment B		Placebo	COMTIs	MAOBIs	Dopamine agonists

Error! No text of specified style in document.

Treatment A				
Placebo	N/A			
COMTIs	-0.81 (-1.01, -0.60)	N/A		
MAOBIs	-0.93 (-1.25, -0.62)	-0.12 (-0.50, 0.25)	N/A	
Dopamine agonists	-1.46 (-1.69, -1.23)	-0.65 (-0.96, -0.35)	-0.53 (-0.92, -0.14)	N/A

UPDRS II (ADL) – RE model



Quantifying heterogeneity/inconsistency:

tau² = 0.2352; l² = 50.9%

Test of heterogeneity/inconsistency:

Q d.f. p.value

24.45 12 0.0176

Differences between treatments – mean and 95% confidence interval

	Treatment A					
Treatment B		Placebo	COMTIs	Dopamine agonists	Amantadine	MAOBIs
	Placebo	N/A				
	COMTIs	-1.47 (-2.12, -0.81)	N/A			
	Dopamine agonists	-1.62 (-2.05, -1.19)	-0.15 (-0.85, 0.54)	N/A		
	Amantadine	-1.70 (-9.11, 5.71)	-0.23 (-7.67, 7.20)	-0.08 (-7.50, 7.34)	N/A	

Error! No text of specified style in document.

Treatment A					
MAOBIs	-1.85	-0.38	-0.23	-0.15	N/A
	(-3.07, -0.63)	(-1.77, 1.00)	(-1.52, 1.06)	(-7.66, 7.36)	

UPDRS III (motor) – RE model



Quantifying heterogeneity/inconsistency:

tau² = 1.2468; l² = 58.2%

Test of heterogeneity/inconsistency:

Q d.f. p.value

28.71 12 0.0044

Differences between treatments – mean and 95% confidence interval

	Treatment A								
Treatment B		Placebo	Amantadine	MAOBIs	COMTIs	Dopamine agonists			
	Placebo	N/A							
	Amantadine	-2.40 (-9.73, 4.93)	N/A						
	MAOBIs	-2.43 (-4.18, -0.68)	-0.03 (-7.56, 7.50)	N/A					
	COMTIs	-3.00	-0.60	-0.57	N/A				

Error! No text of specified style in document.

Treatment A								
	(4.56, -1.44)	(-8.09, 6.89)	(-2.91, 1.77)					
Dopamine agonists	-3.96 (-4.94, -2.99)	-1.56 (-8.95, 5.83)	-1.53 (-3.53, 0.47)	-0.96 (-2.60, 0.67)	N/A			

PDQ-39 – RE model



Quantifying heterogeneity/inconsistency:

tau² = 4.7260; l² = 65.1%

Test of heterogeneity/inconsistency:

Q d.f. p.value

5.72 2 0.0572

Differences between treatments – mean and 95% confidence interval

	Treatment A								
Treatment B		COMTIs	Placebo	Dopamine agonists					
	COMTIs	N/A							
	Placebo	-2.39 (-8.06, 3.29)	N/A						
	Dopamine agonists	-3.89 (-8.90, 1.13)	-1.50 (-4.81, 1.81)	N/A					

Dyskinesia – RE model



Quantifying heterogeneity/inconsistency:

tau² = 0.1426; l² = 62.1%

Test of heterogeneity/inconsistency:

Q d.f. p.value

58 22 < 0.0001



Quantifying heterogeneity/inconsistency:

tau² = 0.0992; l² = 63.7%

Test of heterogeneity/inconsistency:

- Q d.f. p.value
- 60.58 22 < 0.0001

Differences between treatments – relative risk and 95% confidence interval

	Treatment A				
Treatment B		MAOBIs	Placebo	Dopamine agonists	COMTIs
	MAOBIs	N/A			
	Placebo	1.02 (0.53, 1.95)	N/A		
	Dopamine agonists	2.06 (1.04, 4.08)	2.02 (1.62, 2.52)	N/A	
	COMTIs	2.69 (1.30, 5.57)	2.64 (1.88, 3.69)	1.30 (0.92, 1.85)	N/A

Hallucinations – FE model



Quantifying heterogeneity/inconsistency:

tau² = 0.2206; l² = 40.2%

Test of heterogeneity/inconsistency:

Q d.f. p.value

28.42 17 0.0403



Quantifying heterogeneity/inconsistency:

tau² = 0.1407; l² = 31.9%

Test of heterogeneity/inconsistency:

- Q d.f. p.value
- 26.41 18 0.0907

Differences between treatments – relative risk and 95% confidence interval

	Treatment A								
Treatment B		Placebo	COMTIs	MAOBIs	Dopamine agonists				
	Placebo	N/A							
	COMTIs	1.47 (0.99, 2.17)	N/A						
	MAOBIs	1.92 (0.58, 6.34)	1.31 (0.37, 4.60)	N/A					
	Dopamine agonists	2.54 (1.97, 3.28)	1.73 (1.10, 2.73)	1.33 (0.39, 4.51)	N/A				

Mortality – FE model



Quantifying heterogeneity/inconsistency:

tau^2 < 0.0001; I^2 = 100%

Test of heterogeneity/inconsistency:

Q d.f. p.value

0 0 <0.0001

Error! No text of specified style in document.



Quantifying heterogeneity/inconsistency:

tau^2 < 0.0001; I^2 = 100%

Test of heterogeneity/inconsistency:

- Q d.f. p.value
- 0 0 <0.0001.

Differences between treatments – relative risk and 95% confidence interval

	Treatment A							
Treatment B		COMTIs	Dopamine agonists	Placebo				
	COMTIs	N/A						
	Dopamine agonists	1.15 (0.16, 8.33)	N/A					
	Placebo	2.47 (0.56, 10.92)	2.15 (0.58, 7.98)	N/A				

Serious adverse events – FE model



Quantifying heterogeneity/inconsistency:

tau^2 < 0.0001; l^2 = 0%

Test of heterogeneity/inconsistency:

Q d.f. p.value

5.75 8 0.675



Quantifying heterogeneity/inconsistency:

tau^2 < 0.0001; l^2 = 0%

Test of heterogeneity/inconsistency:

- Q d.f. p.value
- 8.03 11 0.7104

Differences between treatments – relative risk and 95% confidence interval

	Treatment A								
Treatment B		Dopamine agonists	Placebo	COMTIs	MAOBIs				
	Dopamine agonists	N/A							
	Placebo	1.15 (0.78, 1.68)	N/A						
	COMTIs	1.25 (0.58, 2.69)	1.09 (0.52, 2.25)	N/A					
	MAOBIs	2.86 (0.53, 15.47)	2.49 (0.48, 12.90)	2.29 (0.38, 13.85)	N/A				

Any adverse event – FE model



Quantifying heterogeneity/inconsistency:

tau² = 0.0028; l² = 31.2%

Test of heterogeneity/inconsistency:

Q d.f. p.value

26.16 18 0.0961



Quantifying heterogeneity/inconsistency:

tau² = 0.0002; l² = 3.6%

Test of heterogeneity/inconsistency:

Q d.f. p.value

20.75 20 0.412

Differences between treatments – relative risk and 95% confidence interval

	Treatment A									
Treatment B		Placebo	MAOBIs	Dopamine agonists	COMTIs					
	Placebo	N/A								
	MAOBIs	1.05 (0.94, 1.17)	N/A							
	Dopamine agonists	1.11 (1.07, 1.14)	1.05 (0.94, 1.18)	N/A						
	COMTIs	1.23 (1.17, 1.31)	1.17 (1.04, 1.33)	1.12 (1.05, 1.19)	N/A					

Adverse event discontinuations – FE model



Quantifying heterogeneity/inconsistency:

tau^2 < 0.0001; l^2 = 0%

Test of heterogeneity/inconsistency:

Q d.f. p.value

17.85 20 0.597



Quantifying heterogeneity/inconsistency:

tau² = 0.0444; l² = 27.4%

Test of heterogeneity/inconsistency:

- Q d.f. p.value
- 30.3 22 0.1114

Differences between treatments – relative risk and 95% confidence interval

	Treatment A								
Treatment B		MAOBIs	Placebo	Dopamine agonists	COMTIs				
	MAOBIs	N/A							
	Placebo	1.43 (0.73, 2.80)	N/A						
	Dopamine agonists	1.47 (0.74, 2.93)	1.03 (0.88, 1.20)	N/A					
	COMTIs	1.84 (0.91, 3.72)	1.28 (1.03, 1.60)	1.25 (0.97, 1.62)	N/A				

E.3 Pharmacological management of non-motor symptoms

E.3.1 Daytime hypersomnolence

Effectiveness of modafinil compared to placebo to treat daytime hypersomnolence

Quality asse	ssment					Number of patients		Effect:mean difference (MD)		
Number of studies	Design	Risk of bias	Inconsiste ncy	Indirectnes s	Imprecisio n	modafinil	placebo	Effect size (95% CI)	Quality	
Epworth slee	Epworth sleepiness scale (ESS)									
4 studies: Ondo (2008) Lou (2009) Hogl (2003) Adler (2002)	RCT	Serious ¹	Serious ²	Not serious	Not serious	53	51	MD -2.01 (-3.08, -0.94)	LOW	
4 studies: Ondo (2008) Lou (2009) Hogl (2003) Adler (2002)	RCT	Serious ¹	Not serious	Not serious	Serious ³	45	46	RR 1.55 (0.99, 2.39)	LOW	
¹ Serious risk	of bias as asse	ssed by NICE	RCT quality ch	ecklist; ² Consi	derable betwee	en study hetero	ogeneity (i ² >40	%); ³ Non-significant result		

E.3.2 Nocturnal akinesia

Quality assess	ment					Number of patients		Effect	
Number of studies	Design	Risk of bias	Inconsiste ncy	Indirectnes s	Imprecisio n	Rotigotine	placebo	Mean difference (95% CI)	Quality
Effect of Rotigo	otine on UPDF	RS-III motor so	core						
Trenkwalder 2010	RCT	Not serious ¹	N/A ²	Not serious ³	Serious ⁴	166	80	-3.55 (-5.37 to -1.73)	MOD
Effect of Rotigo	otine on sleep	quality (PDSS	6 II total score)					
Trenkwalder 2010	RCT	Not serious ¹	N/A ²	Not serious ³	Not serious	166	80	-4.26 (-6.08 to -2.45)	HIGH
Effect of Rotigo	otine on noctu	ırnal akinesia,	dystonia, and	d cramps (NAD	DCS total scor	e)			
Trenkwalder 2010	RCT	Not serious ¹	N/A ²	Not serious ³	Not serious	166	80	-0.41 (-0.79 to -0.04)	HIGH
Effect of Rotigo	otine on numb	per of nocturia	IS						
Trenkwalder 2010	RCT	Not serious ¹	N/A ²	Not serious ³	Serious ⁵	166	80	-0.02 (-0.29 to 0.25)	MOD
Effect of Rotigo	otine on non-r	notor sympto	ms (NMS scal	le)					
Trenkwalder 2010	RCT	Not serious ¹	N/A ²	Not serious ³	Not serious	166	80	-6.65 (-11.99 to -1.31)	HIGH
Effect of Rotigo	otine on activi	ities of daily li	fe (UPDRS -II)						
Trenkwalder 2010	RCT	Not serious ¹	N/A ²	Not serious ³	Not serious ⁶	166	80	-1.49 (-2.32 to -0.65)	HIGH
Effect of Rotigotine on health-related quality of life (PDQ-8)									
Trenkwalder 2010	RCT	Not serious ¹	N/A ²	Not serious ³	Not serious ⁷	166	80	-5.74 (-8.74 to -2.75)	HIGH

¹Low risk of bias as assessed by NICE RCT quality checklist; ²N/A: Not applicable, only 1 study contributed to this analysis; ³No serious indirectness, population as was as specified in the review protocol; ⁴CI cross MID: between 3.25 (Horváth et al., 2015) and 5 points (Schrag et al., 2006); ⁵Non-significant results; ⁶CI do not cross MID of 3 points (Schrag et al., 2006); ⁷CI do not cross MID of 1.6 points (Peto et al., 2001)

Rotigotine effects on early morning motor function and sleep in Parkinson's disease

Adverse events

Quality assessment					Number of patients		Effect		
Number of studies	Design	Risk of bias	Inconsiste ncy	Indirectnes s	Imprecisio n	Rotigotine	Placebo	Risk ratio (95%Cl)	Quality
Adverse events: Rotigotine vs. placebo									
Trenkwalder 2010	RCT	Not serious ¹	N/A ²	Not serious ³	Not serious	166	80	1.27 (1.04 to 1.55)	HIGH
¹ Low risk of bias was as specified	¹ Low risk of bias as assessed by NICE RCT quality checklist; ² N/A: Not applicable, only 1 study contributed to this analysis; ³ No serious indirectness, population as was as specified in the review protocol								

Standard-release compared with controlled-release co-beneldopa

Quality assessment		Effect (nu events)							
Number of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecisio n	Standar d Madopa r	Madopar CR	Quality	
Adverse events									
Madopar Study Group 1989	RCT	Not serious ¹	N/A ²	Not serious ³	Not serious	31	32	High	
¹ Low risk of bias as assessed by NICE RCT quality checklist; ² N/A: Not applicable, only 1 study contributed to this analysis; ³ No serious indirectness, population as was as specified in the review protocol									

Quality assessment								
Number of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Effect	Quality	
Nocturnal disability								
Madopar Study Group 1989	RCT	Not serious ¹	N/A ²	Not serious ³	Serious ⁴	No significant difference ⁴	Moderate	
Early morning disability								

Quality assessment							
Number of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Effect	Quality
Madopar Study Group 1989	RCT	Not serious ¹	N/A ²	Not serious ³	Serious ⁴	No significant difference ⁴	Moderate

¹Low risk of bias as assessed by NICE RCT quality checklist; ²N/A: Not applicable, only 1 study contributed to this analysis; ³No serious indirectness, population as was as specified in the review protocol; ⁴Study reported the results to be non-significant. No numerical data was provided to confirm.

E.3.3 Orthostatic hypotension

Droxidopa for Orthostatic Hypotension

Adverse events

Quality assess	Quality assessment							Effect	
Number of studies	Design	Risk of bias	Inconsiste ncy	Indirectnes s	Imprecisio n	tmt	control	Odds Ratio (95% Cl)	Quality
Total number of adverse events									
2 studies: Hauser 2014 Hauser 2015	RCT	Serious ¹	Serious ²	Not serious	Serious ³	111	111	0.99 (0.51, 1.94)	Very low
¹ Serious risk of bias as assessed by NICE RCT quality checklist; ² Serious inconsistency: I ² = 40% (Cochrane handbook); ³ Non-significant results									

Falls and Fall-related injuries

Quality assess	Quality assessment							Effect	
Number of studies	Design	Risk of bias	Inconsiste ncy	Indirectnes s	Imprecisio n	tmt	control	 Mean Difference (95% Cl) Odds Ratio (95% Cl) 	Quality
Total number o	of patients exp	eriencing fall	related AEs						
2 studies: Hauser 2014 Hauser 2015	RCT	Serious ¹	Not serious	Not serious	Serious ²	111	111	0.56 (0.29, 1.07)	Low

¹Serious risk of bias as assessed by NICE RCT quality checklist; ²Non-significant results

OHQ composite decrease

Quality assess	ment					Number of patients Effect		Effect			
Number of studies	Design	Risk of bias	Inconsiste ncy	Indirectnes s	Imprecisio n	tmt	control	Mean Difference (95% CI)	Quality		
Week 1											
2 studies: Hauser 2014 Hauser 2015	RCT	Serious ¹	Not serious	Not serious	Not serious	111	111	-0.88 (-1.65, -0.11)	Moderate		
Week 2											
2 studies: Hauser 2014 Hauser 2015	RCT	Serious ¹	Not serious	Not serious	Serious ²	111	111	-0.52 (-1.09, 0.05)	Low		
Week 8											
2 studies: Hauser 2014 Hauser 2015	RCT	Serious ¹	Not serious	Not serious	Serious ²	111	111	-0.18 (-0.78, 0.42)	Low		
¹ Serious risk of	Serious risk of bias as assessed by NICE RCT quality checklist: ² Non-significant results										

Mean change in Standing Systolic BP

Quality assess	Quality assessment							Number of patients Effect		
Number of studies	Design	Risk of bias	Inconsiste ncy	Indirectnes s	Imprecisio n	tmt	control	Mean Difference (95% CI)	Quality	
Week 1										
2 studies: Hauser 2014 Hauser 2015	RCT	Serious ¹	Not serious	Not serious	Not serious	111	111	7.34 (2.23, 12.44)	Moderate	
Week 8										

Quality assess	Quality assessment						atients	Effect		
Number of studies	Design	Risk of bias	Inconsiste ncy	Indirectnes s	Imprecisio n	tmt	control	Mean Difference (95% CI)	Quality	
2 studies: Hauser 2014 Hauser 2015	RCT	Serious ¹	Not serious	Not serious	Serious ²	111	111	3.16 (-1.80, 8.12)	Low	
¹ Serious risk of bias as assessed by NICE RCT quality checklist; ² Non-significant results										

Domperidone vs. Fludrocortisone for Orthostatic Hypotension

Adverse events											
Quality assess	Quality assessment						atients	Effect			
Number of studies	Design	Risk of bias	Inconsiste ncy	Indirectnes s	Imprecisio n	domperido ne	fludrocorti sone	Odds Ratio (95% CI)	Quality		
Patients record	ding Adverse	Events									
1 study: Schoffer 2007	RCT	Very Serious ¹	N/A ²	Not serious	Serious ³	13	13	0.73 (0.15, 3.47)	Very Low		
¹ Very serious ri	¹ Very serious risk of bias as assessed by NICE RCT quality checklist; ² N/A: only 1 study contributed to the analysis; ³ Non-significant results										

Blood pressure

Quality assess	Quality assessment							Effect		
Number of studies	Design	Risk of bias	Inconsiste ncy	Indirectnes s	Imprecisio n	domperido ne	fludrocorti sone	Mean Difference (95% CI)	Quality	
Supine blood pressure: mm/Hg										
1 study: Schoffer 2007	RCT	Very Serious ¹	N/A ²	Not serious	Serious ³	13	13	-4 (-23.6 to 15.64)	Very Low	
¹ Very serious	¹ Very serious risk of bias as assessed by NICE RCT quality checklist; ² N/A: only 1 study contributed to the analysis; ³ Non-significant results									

Autonomic function

Quality assess	Quality assessment							Effect	
Number of studies	Design	Risk of bias	Inconsiste ncy	Indirectnes s	Imprecisio n	domperido ne	fludrocorti sone	Mean Difference (95% CI)	Quality
COMPASS:OD									
1 study: Schoffer 2007	RCT	Very Serious ¹	N/A ²	Not serious	Serious ³	13	13	-1 (-2.96 to 0.96)	Very Low
¹ Very serious risk of bias as assessed by NICE RCT quality checklist; ² N/A: only 1 study contributed to the analysis; ³ Non-significant results									

E.3.4 Psychotic symptoms (hallucinations and delusions)

GRADE profile for network meta-analyses

UPDRS Motor

Quality assessment											
Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Quality						
Change in UPDRS motor score											
8	Serious ¹	Not serious	Not serious	Serious ²	LOW						
1 Downgrade 1 level: Limitations in the design or execution of the study											
2 Downgrade 1 level: no interventions had a median rank of 1 [1 to ± n/3]											

BPRS total

Quality assessment					
Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Quality
Change in hallucination score					
7	Serious ¹	Not serious ²	Not serious ³	Not serious	MODERAT E
1 Downgrade 1 level: Limitations in the design or execution of the study 2 Assessed based on residual deviance, deviance information criterion and tau2 (tau2<0.5)					
3 Considered not serious as population, interventions, comparator and outcomes are as defined in protocol					

BPRS Hallucination

Quality assessment					
Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Quality
Change in hallucination score					
3	Serious ¹	Not serious ²	Not serious ³	Not serious	MODERAT E
1 Downgrade 1 level: Limitations in the design or execution of the study 2 Assessed based on residual deviance, deviance information criterion and tau2 (tau2<0.5) 3 Considered not serious as population, interventions, comparator and outcomes are as defined in protocol					

Hallucination – BPRS, NPI, Baylor PD Hallucination, Structured interview for hallucinations in PD

Quality assessment					
Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Quality
Change in hallucination score					
5	Serious ¹	Not serious ²	Not serious ³	Serious ⁴	LOW
 1 Downgrade 1 level: Limitations in the design or execution of the study 2 Assessed based on residual deviance, deviance information criterion and tau2 (tau2<0.5) 3 Considered not serious as population, interventions, comparator and outcomes are as defined in protocol 4 Downgrade 1 level: no interventions had a median rank of 1 [1 to + n(3)] 					

Positive symptoms – SAPS, Positive PANSS, BPRS Positive

Quality assessment					
Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Quality
Change in positive symptom score					
4	Serious ¹	Not serious ²	Not serious ³	Not serious	MODERAT E
1 Downgrade 1 level: Limitations in the design or execution of the study 2 Assessed based on residual deviance, deviance information criterion and tau2 (tau2<0.5)					

3 Considered not serious as population, interventions, comparator and outcomes are as defined in protocol

Treatment discontinuation due to adverse events

Quality assessment					
Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Quality
The rate of an adverse event occurring					
8	Serious ¹	Not serious ²	Not serious ³	Serious ⁴	LOW
1 Downgrade 1 level: Limitations in the design or execution of the study					
2 Assessed based on residual deviance, deviance information criterion and tau2 (tau2<0.5)					
3 Considered not serious as population, interventions, comparator and outcomes are as defined in protocol					

4 Downgrade 1 level: no interventions had a median rank of 1 [1 to \pm n/3]

Adverse events – Estimate of rate

Quality assessment					
Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Quality
Adverse events (Ratio)					

Quality assessment					
Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Quality
5	Serious ¹	Not serious ²	Not serious ³	Not serious ⁴	LOW
1 Downgrade 1 level: Limitations in the design or execution of the study					
2 Assessed based on residual deviance, deviance information criterion and tau2 (tau2<0.5)					
3 Considered not serious as population, in	terventions, comparator and	d outcomes are as defined in pro	otocol		

4 Downgrade 1 level: no interventions had a median rank of 1 [1 to \pm n/3]

Network meta-analyses

Adverse events (rate)



Adverse events (rate) – evidence network

Adverse events (rate) – input data

	Placebo	Clozapine	Olanzapine	Quetiapine	
Morgante et al. (2004) - 0.23yr		5/1722		3/1701	
Ondo et al. (2002) - 0.17yr	12/735		17/1029		
Fernandez et al. (2009) - 0.19yr	11/538.125			9/430.5	
Ondo et al. (2005) - 0.23yr	14/756			23/1596	
Nichols et al. (2013) - 0.08yr	5/224		15/280		
Rate data: numerators represent numbers of AEs; denominators are patient-days of exposure					

Adverse events (rate) - relative effectiveness of all pairwise combinations

	Placebo	Clozapine	Olanzapine	Quetia pine
Placebo		N/A	N/A	N/A
Clozapine	1.55 (0.31, 9.31)		N/A	N/A
Olanzapine	1.43 (0.82, 2.69)	0.92 (0.14, 5.29)		N/A
Quetiapine	0.86 (0.50, 1.53)	0.57 (0.10, 2.58)	0.60 (0.27, 1.35)	

Values given are hazard ratios.

The segment below and to the left of the shaded diagonal is derived from the network meta-analysis, reflecting direct and indirect evidence of treatment effects (row versus column). The point estimate reflects the mean of the posterior distribution, and numbers in parentheses are 95% credible intervals. Because it is not easily possible to pool dichotomous and rate data and derive analogous estimates of hazard ratios from a single frequentist analysis of direct data only, the segment above and to the right of the shaded diagonal is left blank



Adverse events (rate) – relative effect of all options versus common comparator

Adverse events (rate) – rankings for each comparator

	Probability best	Median rank (95%Cl)
Placebo	0.234	2 (1, 3)
Clozapine	0.201	4 (1, 4)
Olanzapine	0.042	3 (1, 4)
Quetiapine	0.523	1 (1, 3)


Adverse events (rate) – rank probability histograms

Adverse events (rate) – model fit statistics

Residual deviance	Dbar	Dhat	рD	DIC
10.42	51.721	43.711	8.01	59.732
(compared to 10 datapoints)				

Adverse events (rate) – notes

- Count (Poisson; log link); fixed effects
- 50000 burn-ins; 10000 recorded iterations

Treatment discontinuation due to AEs



Treatment discontinuation due to AEs – evidence network

Treatment discontinuation due to AEs – input data

	Placebo	Clozapine	Olanzapine	Quetiapine
Morgante et al. (2004)		3/23		2/22

	Placebo	Clozapine	Olanzapine	Quetiapine
Friedman (1999)	3/30	3/30		
Pollak et al. (2004)	2/28	2/32		
Fernandez et al. (2009)	1/8			4/8
Breier et al. (2002) – Europe	1/28		8/49	
Breier et al. (2002) – USA	1/42		10/41	
Nichols et al. (2013)	0/9		7/14	
Shotbolt et al. (2009)	3/13			3/11

Treatment discontinuation due to AEs - relative effectiveness of all pairwise combinations

	Placebo	Clozapine	Olanzapine	Quetiapine
Placebo		0.94 (0.26, 3.45)	10.14 (2.67, 38.50)	2.40 (0.58, 9.87)
Clozapine	1.33 (0.41, 4.49)		-	0.67 (0.10, 4.43)
Olanzapine	15.70 (4.01, 116.30)	12.25 (1.86, 116.70)		-
Quetiapine	1.74 (0.51, 6.29)	1.32 (0.33, 5.52)	0.11 (0.01, 0.73)	

Values given are odds ratios.

The segment below and to the left of the shaded cells is derived from the network meta-analysis, reflecting direct and indirect evidence of treatment effects (row versus column). The point estimate reflects the mean of the posterior distribution, and numbers in parentheses are 95% credible intervals. The segment above and to the right of the shaded cells gives pooled direct evidence (random-effects pairwise meta-analysis), where available (column versus row). Numbers in parentheses are 95% confidence intervals.



Treatment discontinuation due to AEs – relative effect of all options versus common comparator

	Probability best	Median rank (95%Cl)
Placebo	0.589	1 (1, 3)
Clozapine	0.280	2 (1, 3)
Olanzapine	0.000	4 (4, 4)
Quetiapine	0.132	3 (1, 3)

Treatment discontinuation due to AEs – rankings for each comparator



Treatment discontinuation due to AEs – rank probability histograms

Treatment discontinuation due to AEs – model fit statistics

Residual deviance	Dbar	Dhat	pD	DIC
15.52	56.334	45.307	11.028	67.362
(compared to 16 datapoints)				

Treatment discontinuation due to AEs – notes

- Dichotomous synchronic (binomial; logit link); fixed effects
- 50000 burn-ins; 10000 recorded iterations

UPDRS III (motor) score



UPDRS III (motor) score – evidence network

UPDRS III (motor) score - input data

	Placebo	Clozapine	Olanzapine	Quetiapine
Morgante et al. (2004)		-1.30 (9.30)		1.00 (11.00)
Friedman (1999)	-1.80 (6.00)	-3.60 (9.50)		
Pollak et al. (2004)	-3.00 (8.10)	-3.50 (7.70)		
Fernandez et al. (2009)	2.83 (7.46)			-5.74 (6.84)
Breier et al. (2002) – Europe	-0.30 (5.00)		2.70 (6.00)	
Breier et al. (2002) – USA	-0.20 (4.30)		2.60 (6.00)	
Nichols et al. (2013)	1.00 (12.18)		0.80 (12.86)	
Shotbolt et al. (2009)	1.10 (14.69)			-3.00 (13.47)
Values are mean change from baseline to follow	wup(SD)			

UPDRS III (motor) score - relative effectiveness of all pairwise combinations

	Placebo	Clozapine	Olanzapine	Quetiapine
Placebo		-1.09 (-4.06, 1.88)	2.81 (1.16, 4.46)	-7.32 (-13.28, -1.37)
Clozapine	-1.98 (-4.80, 0.78)		-	2.30 (-4.01, 8.61)
Olanzapine	2.82 (1.17, 4.44)	4.80 (1.62, 8.07)		-
Quetiapine	-3.75 (-8.22, 0.70)	-1.75 (-6.29, 2.74)	-6.58 (-11.32, -1.83)	

Values given are weighted mean differences.

The segment below and to the left of the shaded cells is derived from the network meta-analysis, reflecting direct and indirect evidence of treatment effects (row versus column). The point estimate reflects the mean of the posterior distribution, and numbers in parentheses are 95% credible intervals. The segment above and to the right of the shaded cells gives pooled direct evidence (random-effects pairwise meta-analysis), where available (column versus row). Numbers in parentheses are 95% confidence intervals.



UPDRS III (motor) score – relative effect of all options versus common comparator

UPDRS III (motor) score – rankings for each comparator

	Probability best	Median rank (95%CI)
Placebo	0.009	3 (2, 3)
Clozapine	0.219	2 (1, 3)
Olanzapine	0.000	4 (4, 4)
Quetiapine	0.772	1 (1, 3)



UPDRS III (motor) score - rank probability histograms

UPDRS III (motor) score – model fit statistics

Residual deviance	Dbar	Dhat	pD	DIC
15.25	64.259	53.29	10.969	75.228
(compared to 16 datapoints)				

UPDRS III (motor) score - notes

- Continuous (normal; identity link); fixed effects
- 50000 burn-ins; 10000 recorded iterations

BPRS hallucinations



BPRS hallucinations – evidence network

BPRS hallucinations – input data

	Placebo	Olanzapine	Quetiapine
Fernandez et al. (2009)	-0.04 (0.82)		-1.32 (1.13)
Breier et al. (2002) – Europe	-1.40 (1.50)	-1.00 (1.50)	

	Placebo	Olanzapine	Quetiapine
Breier et al. (2002) – USA	-0.90 (1.40)	-0.70 (1.60)	
Values are mean change from baseline to follow up (SD)			

BPRS hallucinations - relative effectiveness of all pairwise combinations

	Placebo	Olanzapine	Quetiapine
Placebo		0.29 (-0.18, 0.77)	-1.28 (-2.25, -0.31)
Olanzapine	0.29 (-0.19, 0.77)		-
Quetiapine	-1.28 (-2.26, -0.31)	-1.58 (-2.65, -0.48)	

Values given are weighted mean differences.

The segment below and to the left of the shaded cells is derived from the network meta-analysis, reflecting direct and indirect evidence of treatment effects (row versus column). The point estimate reflects the mean of the posterior distribution, and numbers in parentheses are 95% credible intervals. The segment above and to the right of the shaded cells gives pooled direct evidence (random-effects pairwise meta-analysis), where available (column versus row). Numbers in parentheses are 95% confidence intervals.



BPRS hallucinations - relative effect of all options versus common comparator

	Probability best	Median rank (95%Cl)
Placebo	0.005	2 (2, 3)
Olanzapine	0.001	3 (2, 3)
Quetiapine	0.994	1 (1, 1)





BPRS hallucinations – rank probability histograms

BPRS hallucinations – model fit statistics

Residual deviance	Dbar	Dhat	pD	DIC
5.17	0.446	-4.555	5	5.446
(compared to 6 datapoints)				

BPRS hallucinations – notes

• Continuous (normal; identity link); fixed effects

• 50000 burn-ins; 10000 recorded iterations

BPRS total



BPRS total – evidence network

BPRS total - input data

	Placebo	Clozapine	Olanzapine	Quetiapine
Morgante et al. (2004)		-10.70 (3.60)		1.60 (4.20)
Friedman (1999)	-2.60 (6.75)	-9.30 (7.79)		
Fernandez et al. (2009)	-0.28 (7.63)			-1.00 (6.97)
Breier et al. (2002)	-5.50 (8.30)		-4.30 (8.30)	
Breier et al. (2002)	-3.10 (5.90)		-2.70 (8.30)	
Shotbolt et al. (2009)	-2.50 (6.40)			-4.20 (6.10)
Rabey (2007)	-4.50 (2.90)			-4.20 (6.00)

BPRS total – relative effectiveness of all pairwise combinations

	Placebo	Clozapine	Olanzapine	Quetiapine
Placebo		-6.70 (-10.59, -2.81)	0.71 (-1.73, 3.15)	-0.58 (-3.66, 2.50)
Clozapine	-9.78 (-12.52, -7.00)		-	12.30 (9.88, 14.72)
Olanzapine	0.71 (-1.72, 3.19)	10.49 (6.79, 14.19)		-
Quetiapine	1.33 (-1.21, 3.92)	11.12 (8.91, 13.30)	0.62 (-2.90, 4.19)	

Values given are hazard ratios.

The segment below and to the left of the shaded diagonal is derived from the network meta-analysis, reflecting direct and indirect evidence of treatment effects (row versus column). The point estimate reflects the mean of the posterior distribution, and numbers in parentheses are 95% credible intervals. Because it is not easily possible to pool dichotomous and rate data and derive analogous estimates of hazard ratios from a single frequentist analysis of direct data only, the segment above and to the right of the shaded diagonal is left blank



BPRS total – relative effect of all options versus common comparator

BPRS total – rankings for each comparator

	Probability best	Median rank (95%CI)
Placebo	0.000	2 (2, 4)
Clozapine	1.000	1 (1, 1)
Olanzapine	0.000	3 (2, 4)
Quetiapine	0.000	4 (2, 4)



BPRS total – rank probability histograms

BPRS total – model fit statistics

Residual deviance	Dbar	Dhat	pD	DIC
15.3	51.777	41.735	10.042	61.819
(compared to 14 datapoints)				

BPRS total – notes

- Continuous (normal; identity link); fixed effects
- 50000 burn-ins; 10000 recorded iterations

Network meta-analyses (pooling across outcomes)

Hallucinations



Hallucinations (multiple scales pooled) - evidence network

Hallucinations (multiple scales pooled) - input data

Study	Scale	PI eb eb	OI an pi ne	Qu eti api
Ondo et al. (2002)	Bespoke interview	-2.80 (4.18)	-3.50 (5.94)	
Fernandez et al. (2009)	BPRS hallucination	-0.04 (0.82)		-1.32 (1.13)
Breier et al. (2002) – Europe	NPS hallucination	-2.70 (3.60)	-2.70 (3.30)	
Breier et al. (2002) – USA	NPS hallucination	-2.50 (2.70)	-2.10 (4.30)	
Shotbolt et al. (2009)	Baylor PD hallucination	-2.50 (5.11)		-3.30 (2.81)
Values are mean change from baseline to follow	v up (SD)			

Hallucinations (multiple scales pooled) - relative effectiveness of all pairwise combinations

	Placebo	Olanzapine	Quetiapine
Placebo		0.03 (-0.26, 0.32)	-0.58 (-1.23, 0.07)
Olanzapine	0.03 (-0.26, 0.32)		-
Quetiapine	-0.61 (-1.25, 0.04)	-0.65 (-1.34, 0.07)	

Values given are standardised mean differences.

The segment below and to the left of the shaded cells is derived from the network meta-analysis, reflecting direct and indirect evidence of treatment effects (row versus column). The point estimate reflects the mean of the posterior distribution, and numbers in parentheses are 95% credible intervals. The segment above and to the right of the shaded cells gives pooled direct evidence (random-effects pairwise meta-analysis), where available (column versus row). Numbers in parentheses are 95% confidence intervals.



Hallucinations (multiple scales pooled) – relative effect of all options versus common comparator

Hallucinations (multiple scales pooled) – rankings for each comparator

	Probability best	Median rank (95%Cl)
Placebo	0.021	2 (2, 3)
Olanzapine	0.030	3 (1, 3)
Quetiapine	0.949	1 (1, 2)



Hallucinations (multiple scales pooled) - rank probability histograms

Hallucinations (multiple scales pooled) – model fit statistics

Residual deviance	Dbar	Dhat	pD	DIC
5.22	3.703	1.721	1.981	5.684
(compared to 5 datapoints)				

Hallucinations (multiple scales pooled) – notes

- Continuous SMD (normal; identity link); fixed effects
- 50000 burn-ins; 10000 recorded iterations

Positive symptoms



Positive symptoms (multiple scales pooled) – evidence network

i ostave symptoms (maniple searcs pooled)	input data			
Study	Scale	PI eb o	CI oz ne	OI an pi ne
Friedman (1999)	SAPS	-3.80 (9.87)	-11.80 (10.39)	
Pollak et al. (2004)	Positive PANSS	-0.80 (2.80)	-5.60 (3.90)	

Positive symptoms (multiple scales pooled) – input data

Study	Scale	PI eb eb	CI oz ne	OI an pi ne
Breier et al. (2002) – Europe	BPRS Positive	-2.90 (3.40)		-2.30 (4.10)
Breier et al. (2002) – USA	BPRS Positive	-1.60 (3.90)		-1.70 (3.50)
Values are mean change from baseline to follow up	o (SD)			

Positive symptoms (multiple scales pooled) - relative effectiveness of all pairwise combinations

	Placebo	Clozapine	Olanzapine
Placebo		-1.09 (-1.48, -0.69)	0.06 (-0.26, 0.37)
Clozapine	-1.11 (-1.50, -0.71)		-
Olanzapine	0.06 (-0.25, 0.37)	1.16 (0.66, 1.67)	

Values given are standardised mean differences.

The segment below and to the left of the shaded cells is derived from the network meta-analysis, reflecting direct and indirect evidence of treatment effects (row versus column). The point estimate reflects the mean of the posterior distribution, and numbers in parentheses are 95% credible intervals. The segment above and to the right of the shaded cells gives pooled direct evidence (random-effects pairwise meta-analysis), where available (column versus row). Numbers in parentheses are 95% confidence intervals.



Positive symptoms (multiple scales pooled) – relative effect of all options versus common comparator

	Probability best	Median rank (95%Cl)
Placebo	0.000	2 (2, 3)
Clozapine	1.000	1 (1, 1)
Olanzapine	0.000	3 (2, 3)

Positive symptoms (multiple scales pooled) – rankings for each comparator



Positive symptoms (multiple scales pooled) – rank probability histograms

Positive symptoms (multiple scales pooled) – model fit statistics

Residual deviance	Dbar	Dhat	pD	DIC
4.624	1.071	-0.91	1.981	3.053
(compared to 4 datapoints)				

Positive symptoms (multiple scales pooled) - notes

- Continuous SMD (normal; identity link); fixed effects
- 50000 burn-ins; 10000 recorded iterations

BPRS psychosis

	Expe	rimen	tal	Co	ontrol	I		Mean Difference	Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Fixed, 95% Cl	IV, Fixed, 95% Cl
20.1.1 Olanzapine									
Nichols MJ et al., 2013 Subtotal (95% Cl)	7.75	4.97	9 9	8	4.9	9 9	100.0% 100.0 %	-0.25 [-4.81, 4.31] - 0.25 [-4.81, 4.31]	
Heterogeneity: Not applic Test for overall effect: Z =	able 0.11 (P	= 0.91)						
Total (95% CI) Heterogeneity: Not applic Test for overall effect: Z = Test for subgroup differen	able 0.11 (P nces: No	= 0.91 ot appl	9) icable			9	100.0%	-0.25 [-4.81, 4.31]	-4 -2 0 2 4 Experimental Control

BPRS psychosis - Clozapine vs. Quetiapine



BPRS hallucination

	Expe	rimen	tal	C	ontrol			Mean Difference	Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Fixed, 95% Cl	IV, Fixed, 95% Cl
20.1.1 Quetiapine									
Fernandez HH et al., 2009	-1.32	1.13	8	-0.04	0.82	8	19.4%	-1.28 [-2.25, -0.31]	
Subtotal (95% CI)			8			8	19.4%	-1.28 [-2.25, -0.31]	
Heterogeneity: Not applicable	е								
Test for overall effect: Z = 2.5	9 (P = 0.	.010)							
20.1.2 Olanzapine									
Breier A et al., 2002 EU	-1	1.5	49	-1.4	1.5	28	37.4%	0.40 [-0.30, 1.10]	
Breier A et al., 2002 US	-0.7	1.6	41	-0.9	1.4	42	43.3%	0.20 [-0.45, 0.85]	
Subtotal (95% CI)			90			70	80.6%	0.29 [-0.18, 0.77]	
Heterogeneity: Chi ² = 0.17, d	f=1 (P=	= 0.68)	; I Z = 0°	Ж					
Test for overall effect: Z = 1.2	1 (P = 0.	.23)							
Total (95% CI)			98			78	100.0%	-0.01 [-0.44, 0.41]	+
Heterogeneity: Chi ² = 8.35, d	f= 2 (P =	= 0.02)	; I ² = 76	6%					
Test for overall effect: Z = 0.0	5 (P = 0.	96)							-Z -I U I Z Everimental Control
Test for subgroup differences	s: Chi ^z =	8.18,	df = 1 (P = 0.00	04), I² =	: 87.8%	6		Experimental Control

Structured interview for hallucinations in PD



Baylor PD hallucination

	Experi	Experimental Control				Mean Difference	Mean Difference				
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Fixed, 95% Cl	IV, Fixed, 95% CI		
20.1.1 Quetiapine											
Shotbolt P et al., 2009 Subtotal (95% Cl)	8.3	2.9	11 11	9.4	4.9	13 13	100.0% 100.0 %	-1.10 [-4.27, 2.07] - 1.10 [-4.27, 2.07]			
Heterogeneity: Not applic	able										
Test for overall effect: Z =	0.68 (P =	= 0.50))								
Total (95% Cl)			11			13	100.0%	-1.10 [-4.27, 2.07]			
Heterogeneity: Not applic	able a co (n -	- 0 50	- 11						-4 -2 0 2 4		
Test for subgroup differen	0.08 (P = nces: No	= 0.51 t app)) licable						Experimental Control		

NPI hallucination

	Expe	rimen	tal	Co	ontrol	I		Mean Difference	Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Fixed, 95% Cl	IV, Fixed, 95% Cl
20.1.1 Olanzapine									
Breier A et al., 2002 EU	-2.7	3.3	49	-2.7	3.6	28	47.7%	0.00 [-1.62, 1.62]	
Breier A et al., 2002 US Subtotal (95% CI)	-2.1	4.3	41 90	-2.5	2.7	42 70	52.3% 100.0 %	0.40 [-1.15, 1.95] 0.21 [-0.91, 1.33]	
Heterogeneity: Chi² = 0.12 Test for overall effect: Z =	2, df = 1 (0.37 (P =	P = 0. : 0.71)	73); I² =)	= 0%					
Total (95% CI) Heterogeneity: Chi ² = 0.12 Test for overall effect: Z =	2, df = 1 (0.37 (P =	P = 0. : 0.71)	90 73); l² =	:0%		70	100.0 %	0.21 [-0.91, 1.33]	-2 -1 0 1 2 Experimental Control

BPRS positive

	Exper	rimen	tal	Co	ontro	I		Mean Difference	Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Fixed, 95% Cl	IV, Fixed, 95% Cl
20.1.1 Olanzapine									
Breier A et al., 2002 EU	-2.3	4.1	49	-2.9	3.4	28	46.6%	0.60 [-1.10, 2.30]	
Breier A et al., 2002 US	-1.7	3.5	41	-1.6	3.9	42	53.4%	-0.10 [-1.69, 1.49]	
Subtotal (95% CI)			90			70	100.0%	0.23 [-0.94, 1.39]	
Heterogeneity: Chi ² = 0.35	5, df = 1 (P = 0.	56); l² =	= 0%					
Test for overall effect: Z =	0.38 (P =	: 0.70)	I						
Total (95% Cl)			90			70	100.0%	0.23 [-0.94, 1.39]	
Heterogeneity: Chi ² = 0.35	5, df = 1 (P = 0.	56); I ² =	= 0%					
Test for overall effect: Z =	0.38 (P =	: 0.70)	I						Fynerimental Control
Test for subgroup differen	ices: Not	t appli	cable						Experimental Control

PANSS positive



SAPS

	Exp	perimenta	l		Control			Mean Difference	Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Fixed, 95% Cl	IV, Fixed, 95% Cl
20.1.1 Clozapine									
Friedman et al., 1999 Subtotal (95% Cl)	-11.8	10.3923	27 27	-3.8	9.87269	27 27	100.0% 100.0 %	-8.00 [-13.41, -2.59] - 8.00 [-13.41, -2.59]	
Heterogeneity: Not appl Test for overall effect: Z	icable = 2.90 (F	P = 0.004)							
Total (95% CI) Heterogeneity: Not appl Test for overall effect: Z Test for subgroup differe	icable = 2.90 (F ences: N	° = 0.004) Vot applica	27 Ible			27	100.0%	-8.00 [-13.41, -2.59]	-10 -5 0 5 10 Experimental Control

NPI delusions

	Exper	imen	tal	Co	ontrol	I I		Mean Difference	Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Fixed, 95% Cl	IV, Fixed, 95% Cl
20.1.1 Olanzapine									
Breier A et al., 2002 EU	-1.1	3.4	49	-2	2.6	28	56.8%	0.90 [-0.45, 2.25]	
Breier A et al., 2002 US	-0.7	3.3	41	-1.7	3.9	42	43.2%	1.00 [-0.55, 2.55]	
Subtotal (95% CI)			90			70	100.0%	0.94 [-0.08, 1.96]	
Heterogeneity: Chi ^z = 0.01	l, df = 1 (l	P = 0.	92); I ² =	= 0%					
Test for overall effect: Z =	1.81 (P =	0.07)	1						
Total (95% Cl)			90			70	100.0%	0.94 [-0.08, 1.96]	
Heterogeneity: Chi ² = 0.01	l, df = 1 (l	P = 0.	92); l² =	= 0%					
Test for overall effect: Z =	1.81 (P =	0.07)	1						-2 -1 U I Z Evnerimental Control
Test for subgroup differer	nces: Not	appli	cable						Experimental Control

UPDRS motor

	Expe	eriment	tal	C	control			Mean Difference	Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Fixed, 95% Cl	IV, Fixed, 95% CI
20.1.1 Quetiapine									
Fernandez HH et al., 2009	-5.74	6.84	8	2.83	7.46	8	63.3%	-8.57 [-15.58, -1.56]	←
Shotbolt P et al., 2009 Subtotal (95% Cl)	28.2	12.3	11 19	30.1	10.4	13 21	36.7% 100.0 %	-1.90 [-11.11, 7.31] - 6.12 [-11.70, -0.54]	
Heterogeneity: Chi ^z = 1.28, c	#f = 1 (P =	= 0.26);	I ² = 22	%					
Test for overall effect: Z = 2.1	15 (P = 0.	.03)							
20.1.2 Olanzapine									
Breier A et al., 2002 EU	2.7	6	49	-0.3	5	28	43.9%	3.00 [0.50, 5.50]	
Breier A et al., 2002 US	2.6	6	41	-0.2	4.3	42	54.2%	2.80 [0.55, 5.05]	│ ──■ ──
Nichols MJ et al., 2013	30.3	13.39	9	31	13.09	9	1.8%	-0.70 [-12.93, 11.53]	·
Subtotal (95% CI)			99			79	100.0%	2.82 [1.17, 4.48]	-
Heterogeneity: Chi² = 0.34, c	f=2 (P =	= 0.84);	$ ^{2} = 0\%$)					
Test for overall effect: Z = 3.3	34 (P = 0.	.0008)							
20.1.3 Clozapine									
Friedman et al., 1999	-3.6	9.5	25	-1.8	6	25	45.4%	-1.80 [-6.20, 2.60]	_
Pollak P et al., 2004	-3.5	7.7	32	-3	8.1	28	54.6%	-0.50 [-4.51, 3.51]	
Subtotal (95% CI)			57			53	100.0%	-1.09 [-4.06, 1.88]	
Heterogeneity: Chi ² = 0.18, c	∄f = 1 (P =	= 0.67);	$ ^{2} = 0\%$,					
Test for overall effect: Z = 0.7	72 (P = 0.	.47)							
									Experimental Control

Test for subgroup differences: Chi² = 12.52, df = 2 (P = 0.002), l² = 84.0%

UPDRS motor – Clozapine vs. Quetiapine

	Cloz	zapin	e	Que	tiapir	ie		Mean Difference	Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Fixed, 95% Cl	IV, Fixed, 95% Cl
20.4.4 Clozapine vs. Quet	iapine								
Morgante L et al., 2004 Subtotal (95% Cl)	56.7	9.2	20 20	54	11	20 20	100.0% 100.0 %	2.70 [-3.58, 8.98] 2.70 [-3.58, 8.98]	
Heterogeneity: Not applica Test for overall effect: Z = (able 0.84 (P	= 0.4	0)						
Total (95% CI) Heterogeneity: Not applica Test for overall effect: Z = (Test for subgroup differen	able 0.84 (P : ices: No	= 0.4 ot app	20 0) blicable	1		20	100.0%	2.70 [-3.58, 8.98]	-4 -2 0 2 4 Clozapine Quetiapine

Treatment discontinuation due to AEs

	Ехрегіт	ental	Contr	ol		Odds Ratio	Odds Ratio
Study or Subgroup	Events	Total	Events	Total	Weight	M-H, Fixed, 95% Cl	M-H, Fixed, 95% Cl
20.1.1 Quetiapine							
Fernandez HH et al., 2009	4	8	1	8	5.4%	7.00 [0.57, 86.32]	
Ondo WG et al., 2005	0	21	0	10		Not estimable	
Shotbolt P et al., 2009	3	11	3	13	21.5%	1.25 [0.20, 7.96]	
Subtotal (95% CI)		40		31	26.8%	2.40 [0.58, 9.87]	
Total events	7		4				
Heterogeneity: Chi ² = 1.17, c	if = 1 (P = 0).28); <mark>I</mark> ²∶	= 15%				
Test for overall effect: Z = 1.2	21 (P = 0.2)	2)					
20.1.2 Olanzapine							
Breier A et al., 2002 EU	8	49	1	28	11.4%	5.27 [0.62, 44.55]	
Breier A et al., 2002 US	10	41	1	42	8.0%	13.23 [1.61, 108.86]	
Nichols MJ et al., 2013	7	14	0	9	3.2%	19.00 [0.93, 388.77]	
Ondo WG et al., 2002	0	18	0	12		Not estimable	
Subtotal (95% CI)		122		91	22.7%	10.03 [2.64, 38.13]	
Total events	25		2				
Heterogeneity: Chi ² = 0.59, c	f = 2 (P = 0).75); l²∶	= 0%				
Test for overall effect: Z = 3.3	39 (P = 0.00	007)					
20.1.3 Clozapine							
Friedman et al., 1999	3	30	3	30	29.0%	1.00 [0.19, 5.40]	+
Pollak P et al., 2004	2	32	2	28	21.5%	0.87 [0.11, 6.59]	•
Subtotal (95% CI)		62		58	50.5%	0.94 [0.26, 3.45]	
Total events	5		5				
Heterogeneity: Chi ² = 0.01, c	f=1 (P=0).92); I ²∶	= 0%				
Test for overall effect: Z = 0.0)9 (P = 0.93	3)					
Total (95% CI)		224		180	100.0%	3.40 [1.67, 6.91]	
Total events	37		11				
Heterogeneity: Chi ² = 8.21, c	f = 6 (P = 0).22); I ²	= 27%			-	
Test for overall effect: Z = 3.3	38 (P = 0.00	007)					Experimental Control
Test for subgroup difference	s: Chi ^z = 6	.26, df=	: 2 (P = 0	.04), I ^z :	= 68.0%		Experimental Control

Treatment discontinuation due to AEs – Clozapine vs. Quetiapine

	Experime	ental	Contr	ol				Odds	s Ratio			
Study or Subgroup	Events	Total	Events	Total	Weight	M-H, Fixed, 95% Cl			M-H, Fixe	ed, 95% Cl		
20.5.4 Clozapine vs. Que	etiapine											
Morgante L et al., 2004 Subtotal (95% Cl)	3	23 23	2	22 22	100.0% 100.0 %	1.50 [0.23, 9.96] 1.50 [0.23, 9.96]						
Total events Heterogeneity: Not applic Test for overall effect: Z =	3 able : 0.42 (P = 1	0.67)	2									
Total (95% CI)		23		22	100.0%	1.50 [0.23, 9.96]						
Total events Heterogeneity: Not applic Test for overall effect: Z = Test for subgroup differe	3 cable : 0.42 (P = 1 nces: Not a	D.67) applicat	2 Die				⊢ 0.1	0.2	0.5 Clozapine	1 2 Quetiapine	<u> </u> 5	10

Adverse events (rate)

						Rate	Ratio	
Study or Subgroup	log[Rate Ratio]	SE	Weight	IV, Fixed, 95% Cl		IV, Fixed	1, 95% Cl	
20.2.1 Quetiapine								
Ondo WG et al., 2005	-0.25078	0.338979	39.3%	0.78 [0.40, 1.51]				
Fernandez HH et al., 2009 Subtotal (95% CI)	0.022473	0.449467	22.4% 61.7 %	1.02 [0.42, 2.47] 0.86 [0.51, 1.46]				
Heterogeneity: Chi ² = 0.24, Test for overall effect: Z = 0.	df = 1 (P = 0.63); l² 56 (P = 0.58)	= 0%						
20.2.2 Olanzapine								
Nichols MJ et al., 2013	0.875469	0.516398	17.0%	2.40 [0.87, 6.60]		_	•	 →
Ondo WG et al., 2002 Subtotal (95% Cl)	0.962811	0.460566	21.3% 38.3 %	2.62 [1.06, 6.46] 2.52 [1.28, 4.94]				→
Heterogeneity: Chi ² = 0.02,	$df = 1 (P = 0.90); I^2$:	= 0%						
Test for overall effect: Z = 2	.69 (P = 0.007)							
Total (95% CI)			100.0%	1.30 [0.85, 1.97]		-		
Heterogeneity: Chi ² = 6.30,	$df = 3 (P = 0.10); I^2$:	= 52%			\vdash	0.5		
Test for overall effect: Z = 1.	.22 (P = 0.22)				0.2	0.5 Favours (experimental)	Favours (control)	5
Test for subgroup differenc	es: Chi ² = 6.05, df =	: 1 (P = 0.01	l), I ^z = 83.	.5%		, alloard foxportitiontail	, aroaro [control]	

Adverse events (rate) – Clozapine vs. Quetiapine



Mortality

	Experime	ental	Contr	ol		Odds Ratio	Odds Ratio			
Study or Subgroup	Events	Total	Events	Total	Weight	M-H, Fixed, 95% Cl	M-H, Fixed, 95% Cl			
20.1.1 Quetiapine										
Fernandez HH et al., 2009	0	8	0	8		Not estimable				
Ondo WG et al., 2005	0	21	2	10	65.2%	0.08 [0.00, 1.82]	<			
Shotbolt P et al., 2009	0	11	0	13		Not estimable				
Subtotal (95% Cl)		40		31	65.2%	0.08 [0.00, 1.82]				
Total events	0		2							
Heterogeneity: Not applicable	е									
Test for overall effect: Z = 1.5	8 (P = 0.11)								
20.1.2 Olanzapine										
Nichols MJ et al., 2013	0	14	1	9	34.8%	0.20 (0.01, 5.35)	_			
Ondo WG et al., 2002	0	18	0	12		Not estimable				
Subtotal (95% Cl)		32		21	34.8%	0.20 [0.01, 5.35]				
Total events	0		1							
Heterogeneity: Not applicable	е									
Test for overall effect: Z = 0.9	7 (P = 0.33	3)								
Total (95% CI)		72		52	100.0%	0.12 [0.01, 1.14]				
Total events	Ο		3							
Heterogeneity: $Chi^2 = 0.15$ d	f=1(Ρ=0		= 0%				+ + + +			
Test for overall effect: $7 = 1.8$	5 (P = 0.00	3) 11	- 0 /0				0.005 0.1 1 10			
Test for subgroup differences	s:Chi ² = 0.00	″ 15 df=	:1 (P = 0	70) P:	= 0%		Experimental Control			
reactor aubgroup unerences	5. Om = 0.	10, ai -	. (1 = 0		0.0					

BPRS Psychosis - Olanzapine

Quality assessment	:		Number of p	atients	Effect					
Number of studies	Desig n	Risk of bias	Inconsistenc y	Indirectnes s	Imprecisio n	Interventio n	Contr ol	Mean Difference (95% CI)	Qualit y	
Average CI score change										
1 study: Nichols et al., 2013	RCT	Serious ¹	N/A ²	Not serious ³	Serious ⁴	9	9	-0.25 (-4.81, 4.31)	LOW	
Nichols et al., 2013 ¹ Serious risk of bias as assessed by NICE RCT quality checklist ² N/A; Not applicable, only 1 study contributed to this analysis ³ No serious indirectness, population was as specified in review protocol ⁴ Non-significant regults										

BPRS Psychosis - Clozapine vs. Quetiapine

Quality assessment			Number of p	atients	Effect	Qualit			
Number of studies	Desig	Risk of	Inconsistenc	Indirectnes	Imprecisio n	Interventio n	Contr	Mean Difference (95% CI)	У
Average Cl score char		0103	У	3			UI	Mean Difference (55 % Ci)	1
Average CI score char	iye								
1 study:	RCT	Serious ¹	N/A ²	Not serious ³	serious ⁴	20	20	0.1 (-1, 1.2)	
									LOW
Morgante et al., 2004									
¹ Serious risk of bias as asse	essed by N	CE RCT quality cl	necklist						
² N/A; Not applicable, only 1	study contr	ibuted to this anal	ysis						
³ No serious indirectness, po	pulation wa	as as specified in r	review protocol						
⁴ Non-significant results									

BPRS Hallucination – Quetiapine

Quality assessment			Number of p	atients	Effect							
Number of studies	Desig n	Risk of bias	Inconsistenc v	Indirectnes s	Imprecisio n	Interventio n	Contr ol	Mean Difference (95% CI)	Qualit v			
Average CI score chang	je		-									
1 study:	RCT	Serious ¹	N/A ²	Not serious ³	Serious ⁶	8	8	-1.28 (-2.25, -0.31)	LOW			
	Quality assessment			Number of p	atients	Effect						
---	--	------------	--------------	-------------------	------------------	-----------------	------------------	-------------	--------------------------	-------------	--	--
	Number of studies	Desig n	Risk of bias	Inconsistenc v	Indirectnes s	Imprecisio n	Interventio n	Contr ol	Mean Difference (95% Cl)	Qualit v		
Fernandez et al., 2009												
¹ Serious risk of bias as assessed by NICE RCT quality checklist ² N/A; Not applicable, only 1 study contributed to this analysis ³ No serious indirectness, population was as specified in review protocol ⁴ Non-significant results												
	⁵ Serious imprecision: CI cross MID between 3.25 (Horvath et al., 2015) and 5 points (Schrag et al., 2006) ⁶ Very small sample size											

BPRS Hallucination – Olanzapine

Quality assessment					Number of patients		Effect					
Number of studies	Desi gn	Risk of bias	Inconsisten cy	Indirectne ss	Imprecisi on	Interventi on	Contr ol	Mean Difference (95% Cl)	Quality			
Average CI score change												
2 studies: RCT Serious ¹ Not serious Not serious ³ Serious ⁴ 90 70 $0.29(-0.18, 0.77)$ LOW												
Breier et al., 2002 – EU study												
Breier et al., 2002 – US study	Breier et al., 2002 – US study											
 Serious risk of bias as assessed by N/A; Not applicable, only 1 study of No serious indirectness, population 4 Non-significant results 	Study ¹ Serious risk of bias as assessed by NICE RCT quality checklist ² NA; Not applicable, only 1 study contributed to this analysis ³ No serious indirectness, population was as specified in review protocol ⁴ Non-significant results											

NPI hallucination – Olanzapine

Quality assessment						Number of patients		Effect	Qual
Number of studies	Desi gn	Risk of bias	Inconsiste ncy	Indirectn ess	Imprecisi on	Interventi on	Cont rol	Mean Difference (95% CI)	ity
Average CI score change						'			

Quality assessment					Number of patients		Effect			
Number of studies	Desi gn	Risk of bias	Inconsiste ncy	Indirectn ess	Imprecisi on	Interventi on	Cont rol	Mean Difference (95% CI)	ity	
2 studies: Breier et al., 2002 – EU study Breier et al., 2002 – US study	RCT	Serious ¹	Not serious	Not serious ³	Serious ⁴	90	70	0.21 (-0.91, 1.33)	LOW	
Serious risk of bias as assessed by NICE RCT quality checklist N/A; Not applicable, only 1 study contributed to this analysis No serious indirectness, population was as specified in review protocol Non-significant results										

Baylor PD Hallucination – Quetiapine

Quality assessment				Number of p	atients	Effect				
Number of studies	Desig n	Risk of bias	Inconsistenc y	Indirectnes s	Imprecisio n	Interventio n	Contr ol	Mean Difference (95% CI)	Qualit y	
Average CI score cha	inge									
1 study: Shotbolt et al., 2009	RCT	Serious ¹	N/A ²	Not serious ³	Serious ⁴	11	13	-1.1 (-4.27, 2.07)	LOW	
 ¹ Serious risk of bias as assessed by NICE RCT quality checklist ² N/A; Not applicable, only 1 study contributed to this analysis ³ No serious indirectness, population was as specified in review protocol ⁴ Non-significant results 										

Structured interview for hallucinations in PD – Olanzapine

Quality assessmen	nt					Number of patients		Effect	Quali
Number of	Desi	Risk of	Inconsistenc	Indirectnes	Imprecisio	Interventio	Contr		ty
studies	gn	bias	у	S	n	n	Ol	Mean Difference (95% CI)	
Average CI score ch	nange								
1 study:	RCT	Serious ¹	N/A ²	Not serious ³	Serious ⁴	16	11	-1.6 (-5.94, 2.74)	LOW

Quality assessmer	nt				Number of p	oatients	Effect	Quali			
Number of	Number of Desi Risk of Inconsistenc Indirectnes Imp								ty		
studies	gn	bias	У	n	n	ol	Mean Difference (95% CI)				
Ondo et al., 2002											
¹ Serious risk of bias as a	assessed b	y NICE RCT qual	ity checklist								
² N/A; Not applicable, onl	N/A; Not applicable, only 1 study contributed to this analysis										
³ No serious indirectness	No serious indirectness, population was as specified in review protocol										

⁴ Non-significant results

BPRS Positive – Olanzapine

Quality assessment				No of patier	nts	Effect	Quali		
No of studies	Desi gn	Risk of bias	Inconsistenc y	Indirectne ss	Imprecisio n	Interventio n	Contr ol	Mean Difference (95% Cl)	ty
Average CI score change					·		·		
2 studies:	RCT	Serious ¹	Not serious	Not serious ³	Serious ⁴	90	70	0.23 (-0.94, 1.39)	LOW
Breier et al., 2002 – EU study									
Breier et al., 2002 – US study									
 Serious risk of bias as assessed by N/A; Not applicable, only 1 study cor No serious indirectness, population Non-significant results 	NICE RCT ntributed to was as spe	quality checklist this analysis ecified in review p	rotocol						

Positive PANSS – Clozapine

Quality assessmen	it				Number of p	patients	Effect				
Number of studies	Desig n	Risk of bias	Inconsistenc y	Indirectnes s	Imprecisio n	Interventio n	Contr ol	Mean Difference (95% CI)	Quality		
Average CI score change											
1 study:	RCT	Serious ¹	N/A ²	Not serious ³	Not serious	32	28	-4.8 (-6.5, -3.1)	MODERAT F		
Pollak et al., 2004									-		
 ¹ Serious risk of bias as assessed by NICE RCT quality checklist ² N/A; Not applicable, only 1 study contributed to this analysis ³ No serious indirectness, population was as specified in review protocol 											

SAPS – Clozapine

Quality assessment					Number of p	atients	Effect	Qualit			
Number of studies	Desig n	Risk of bias	Inconsistenc y	Indirectnes s	Imprecisio n	Interventio n	Contr ol	Mean Difference (95% CI)	У		
Average CI score change											
1 study: Friedman et al., 1999	RCT	Serious ¹	N/A ²	Not serious ³	Serious⁴	27	27	-8 (-13.41, - 2.59)	LOW		
 ¹ Serious risk of bias as assessed by NICE RCT quality checklist ² N/A; Not applicable, only 1 study contributed to this analysis ³ No serious indirectness, population was as specified in review protocol ⁴ Non-significant results 											

NPI Delusions – Olanzapine

Quality assessment				Number of patients		Effect	1		
Number of studies	Desi gn	Risk of bias	Inconsistenc y	Indirectne ss	Imprecisio n	Interventio n	Contr ol	Mean Difference (95% Cl)	Quali ty
Average CI score change									
2 studies: Breier et al., 2002 – EU study	RCT	Serious ¹	Not serious	Not serious ³	Serious ⁴	90	70	0.94 (-0.08, 1.96)	LOW
Breier et al., 2002 – US study									
 Serious risk of bias as assessed by N/A; Not applicable, only 1 study co No serious indirectness, population 	NICE RCT ntributed to was as spe	quality checklist this analysis ecified in review p	protocol						

⁴ Non-significant results

UPDRS Motor – Quetiapine

Error! No text of specified style in document.

Quality assessment				Number of pa	atients	Effect	Qualit				
Number of studies	Desig n	Risk of bias	Inconsistency	Indirectness	Imprecision	Intervention	Contr ol	Mean Difference (95% Cl)	У		
JPDRS Motor - Quetiapine (Better indicated by lower values)											
2 studies: Fernandez et al., 2009 Shotbolt et al., 2009	RCT	Serious ¹	Not serious	Not serious ³	Serious ⁵	19	21	-6.12 (-11.7, - 0.54)	LOW		
 Should be et al., 2003 Serious risk of bias as assessed by NICE RCT quality checklist ² NA; Not applicable, only 1 study contributed to this analysis ³ No serious indirectness, population was as specified in review protocol ⁴ Non-significant results ⁵ Serious imprecision: Cl cross MID between 3.25 (Horvath et al., 2015) and 5 points (Schrag et al., 2006) 											

UPDRS Motor - Olanzapine

Quality assessment				Number of patients		Effect	1		
Number of studies	Desi gn	Risk of bias	Inconsisten cy	Indirectne ss	Imprecisi on	Interventi on	Contr ol	Mean Difference (95% Cl)	Quali ty
Average CI score change									
3 studies:	RCT	Serious ¹	Not serious	Not serious ³	Serious ⁵	99	79	2.82 (1.17, 4.48)	LOW
Breier A et al., 2002 - EU study									
Breier A et al., 2002 – US study									
Nichols et al., 2013									
 Serious risk of bias as assessed by N N/A; Not applicable, only 1 study controls No serious indirectness, population wa Non-significant results Serious imprecision: CI cross MID bet 	ICE RCT of ributed to t as as spec ween 3.25	uality checklist his analysis ified in review pro (Horvath et al., 2	otocol 2015) and 5 points (\$	Schrag et al., 2006	6)				

UPDRS Motor – Clozapine

Quality assessment						Number of p	atients	Effect	Qualit
Number of studies	Desig n	Risk of bias	Inconsistenc y	Indirectnes s	Imprecisio n	Interventio n	Contr ol	Mean Difference (95% CI)	У
Average CI score char	ige								
2 studies:	RCT	Serious ¹	Not serious	Not serious ³	Serious ⁵	57	53	-1.09 (-4.06, 1.88)	LOW
Friedman et al., 1999									
Pollak et al., 2004									
 ¹ Serious risk of bias as asse ² N/A; Not applicable, only 1 ³ No serious indirectness, po ⁴ Non-significant results ⁵ Serious imprecision: Clara 	essed by NI study contropulation wa	CE RCT quality ch ibuted to this analy as as specified in r ween 3 25 (Horvat	ecklist /sis eview protocol h et al. 2015) and 5 r	points (Schrag et al	2006)				

UPDRS Motor - Clozapine vs. Quetiapine

Quality assessment						Number of p	atients	Effect	
Number of studies	Desig n	Risk of bias	Inconsistenc y	Indirectnes s	Imprecisio n	Interventio n	Contr ol	Mean Difference (95% CI)	Qualit y
Average CI score char	ige								'
1 study:	RCT	Serious ¹	N/A ²	Not serious ³	Serious ⁵	20	20	2.7 (-3.58, 8.98)	LOW
Morgante et al., 2004									

Dropouts due to AEs – Quetiapine

Quality assessment						No of events/ Total n	o of patients	Effect	
Number of studies	Desig n	Risk of bias	Inconsistenc y	Indirectnes s	Imprecisio n	Intervention	Control	Odds Ratio (95% Cl)	Quali ty
Dropouts due to AEs									
3 studies:	RCT	Serious ¹	Not serious	Not serious ³	Serious ⁴	7/40	4/31	2.4 (0.58, 9.87)	LOW
Fernandez et al., 2009									
Ondo et al., 2005 Shotbolt et al., 2005									

Quality assessment						No of events/ Total n	o of patients	Effect	
Desig Risk of Inconsistenc Indirectnes Imp Number of studies n bias v s n								Odds Ratio	Quali
Number of studies n bias y s n						Intervention	Control	(95% CI)	ty
¹ Serious risk of bias as asses	ssed by NI	CE RCT quality ch	necklist						
² N/A; Not applicable, only 1 s	study contri	buted to this analy	ysis						
³ No serious indirectness, pop	oulation wa	s as specified in r	eview protocol						
⁴ Non-significant results									

Dropouts due to AEs – Olanzapine

Quality assessment						No of events/ Total patients	no of	Effect	
No of studies	Desi gn	Risk of bias	Inconsisten cy	Indirectne ss	Imprecisi on	Intervention	Control	Odds Ratio (95% Cl)	Quality
Dropouts due to AEs									
4 studies:	RCT	Serious ¹	Not serious	Not serious ³	Not serious	25/122	2/91	10.03 (2.64, 38.13)	MODERA TE
Breier et al., 2002 – EU									
Breier et al., 2002 – US									
Nichols et al., 2013 Ondo et al., 2002									
 ¹ Serious risk of bias as asse ² N/A; Not applicable, only 1 ³ No serious indirectness, po 	essed by N study con pulation v	NICE RCT quality tributed to this ar vas as specified i	r checklist nalysis in review protocol						

Dropouts due to AEs – Clozapine

Quality assessmen	it					No of events/ T patients	otal no of	Effect	
Number of studies	Desi an	Risk of bias	Inconsisten	Indirectne	Imprecisi on	Intervention	Control	Odds Ratio (95% CI)	Qual ity
Dropouts due to AEs	6								

Quality assessmen	nt					No of events/ T patients	otal no of	Effect	
Number of studies	Desi gn	Risk of bias	Inconsisten cy	Indirectne ss	Imprecisi on	Intervention	Control	Odds Ratio (95% CI)	Qual ity
2 studies:	RCT	Serious ¹	Not serious	Not serious ³	Serious ⁴	5/62	5/58	0.94 (0.26 to 3.45)	LOW
Friedman et al., 1999									
Pollak et al., 2014									
 Serious risk of bias as a 2 N/A; Not applicable, only No serious indirectness, 4 Non-significant results 	assessed k y 1 study , populatic	by NICE RCT qui contributed to thi on was as specifi	ality checklist s analysis ed in review protocc	bl					

Dropouts due to AEs - Clozapine vs. Quetiapine

Quality assessmen	t					No of events/ T patients	otal no of	Effect	
Number of studies	Desi gn	Risk of bias	Inconsisten cy	Indirectne ss	Imprecisi on	Intervention	Control	Odds Ratio (95% Cl)	Quali ty
Dropouts due to AEs	3								
1 study:	RCT	Serious ¹	N/A ²	Not serious ³	Serious ⁴	3/23	2/22	1.5 (0.23, 9.96)	LOW
Morgante et al., 2004									
 Serious risk of bias as a N/A; Not applicable, only No serious indirectness, Non-significant results 	ssessed by 1 study c populatior	y NICE RCT qua ontributed to this n was as specifie	lity checklist analysis d in review protocol						

Adverse event - Estimate of rate - Quetiapine

Quality ass	essment					Number of pati	ents	Effect	Qualit
Number of studies	umber Risk of bias Inconsistency Indirectness Imprecision						Control	Rate Ratio (95% Cl)	У
The rate of	an adverse eve	nt occurring							

Error! No text of specified style in document.

Quality ass	essment					Number of pati	ents	Effect	Qualit
Number of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Intervention	Control	Rate Ratio (95% Cl)	У
2 studies: Fernandez et al., 2009 Ondo et al., 2005	RCT	Serious ¹	Not serious	Not serious ³	Serious ⁴	29	18	0.86 (0.51, 1.46)	LOW
 Serious risk c N/A; Not appl No serious ind Non-signification 	of bias as assessed icable, only 1 study directness, populati nt results	by NICE RCT que contributed to the on was as speci	uality checklist nis analysis fied in review protocol						

Adverse event - Estimate of rate - Olanzapine

Quality assessment	t					Number of pa	tients	Effect	í l
Number of studies	Desig n	Risk of bias	Inconsistency	Indirectnes s	Imprecisio n	Intervention	Control	Rate Ratio (95% CI)	Quality
The rate of an advers	se event c	occurring							
2 studies: Nichols et al., 2013	RCT	Serious ¹	Not serious⁵	Not serious ³	Not serious	31	21	2.52 (1.28, 4.94)	MODERAT E
Ondo et al., 2002									
 Serious risk of bias as as N/A; Not applicable, only No serious indirectness, Non-significant results 	ssessed by N 1 study con population v	NICE RCT quality c tributed to this anal vas as specified in	hecklist lysis review protocol						

Adverse event - Estimate of rate - Clozapine vs. Quetiapine

Quality assessment				Number of pa	tients	Effect			
Number of studies	Desig n Risk of bias Inconsistency Indirectness Imprecisio							Rate Ratio (95% CI)	Qualit y
The rate of an adverse ev	vent occur	ring		,					
1 study:	RCT	Serious ¹	N/A ²	Not serious ³	Serious ⁴	23	22	1.65 (0.39, 6.89)	LOW

Quality assessment				Number of par	tients	Effect			
	Desig						Contro	Rate Ratio	Qualit
Number of studies	n	Risk of bias	Inconsistency	Indirectness	Imprecision	Intervention	1	(95% CI)	у
Morgante et al., 2004									
 ¹ Serious risk of bias as assessed by NICE RCT quality checklist ² N/A; Not applicable, only 1 study contributed to this analysis ³ No serious indirectness, population was as specified in review protocol 									

Mortality - Quetiapine

Quality assessment						No of events/ T patients	otal no of	Effect			
Number of studies	Desi gn	Risk of bias	Inconsisten cy	Indirectne ss	Imprecisi on	Intervention	Control	Odds Ratio (95% Cl)	Qualit y		
Mortality											
3 studies:	RCT	Serious ¹	N/A ²	Not serious ³	Serious ⁴	0/40	2/31	OR 0.08 (0, 1.82)	LOW		
Fernandez et al., 2009											
Ondo et al., 2005											
Shotbolt et al., 2009											
Serious risk of bias as assessed by NICE RCT quality checklist ² N/A; Not applicable, only 1 study contributed to this analysis ³ No serious indirectness, population was as specified in review protocol ⁴ Non-significant results											

Mortality – Olanzapine

Quality assessme	ent					No of events/ Total patients	Effect		
Number of studies	Desi gn	Risk of bias	Inconsisten cy	Indirectne ss	Imprecisi on	Intervention	Control	Odds Ratio (95% Cl)	Quali ty
Mortality - Olanzap	ine								
2 studies:	RCT	Serious ¹	N/A ²	Not serious ³	Serious ⁴	0/32	1/21	OR 0.2 (0.01, 5.35)	LOW

Error! No text of specified style in document.

Quality assessme	ent					No of events/ Total no of patients Effect			
Number of studies	lumber of Desi Risk of Inconsisten Indirectne on ss						Control	Odds Ratio (95% CI)	Quali ty
Nichols et al., 2013 Ondo et al., 2002									
 Serious risk of bias as N/A; Not applicable, or No serious indirectnes Non-significant results 	assessed nly 1 study s, populat	by NICE RCT q contributed to the tot of tot o	uality checklist his analysis fied in review protoc	col					

E.3.5 REM sleep disorder behaviour

Rivastigmine effects on RBD sleep disorder in Parkinson's disease

Quality assess	ment					Number of patients Effect			
Number of studies	Design	Risk of bias	Inconsiste ncy	Indirectnes s	Imprecisio n	Rivastigmi ne	placebo	Median difference (25 th - 75 th %ile)	Quality
Frequency of F	RBD episodes								
Di Giacomo 2012	RCT	Serious ¹	NA ²	Not serious	Serious ³	12	12	2.5 (0.0 to 4.5)	LOW
¹ Very serious risk of bias as assessed by NICE RCT quality checklist; ² N/A: only 1 study contributed to the analysis; ³ Study number is very small									

Rivastigmine for the treatment of RBD sleep disorder: Serious adverse events

Quality assess	ment					Number of p	atients	Effect	
Number of studies	Design	Risk of bias	Inconsiste	Indirectnes s	Imprecisio n	Rivastigmi	placebo	Number of adverse events leading to discontinuation	Quality
Adverse events	s leading to s	tudy discontin	nuation in riva	stigmine grou	ıp		placebe		Quanty
Di Giacomo 2012	RCT	Serious ¹	NA ²	12	12	2	LOW		
A .L		4							

Adverse events leading to study discontinuation in placebo group

Quality assess	ment			Number of p	atients	Effect			
Number of studies	Design	Risk of bias	Inconsiste ncy	Indirectnes s	Imprecisio n	Rivastigmi ne	placebo	Number of adverse events leading to discontinuation	Quality
Di Giacomo 2012	RCT	Serious ¹	NA ²	Not serious	Serious ³	12	12	0	LOW

¹Very serious risk of bias as assessed by NICE RCT quality checklist; ²N/A: only 1 study contributed to the analysis; ³Study number is very small

E.3.6 Thermoregulatory dysfunction

None

E.4 Pharmacological management of dementia associated with Parkinson's disease

Parkinson's disease dementia – cholinesterase inhibitors

PDD – cholinesterase inhibitor vs. placebo: adverse events										
Quality assessment No of patients										
No of studios Dosign	Improcision	Chl	Placabo	Polat						

			,							A
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Chl	Placebo	Relative (95% CI)	Absolute (95% Cl)	Quality
Any adverse e	vents –	cholinesteras	e inhibitors (pro	bability of exp	eriencing ≥1;	follow-u	p 10 to 24	weeks; lower is better)	; see Figure 1 for forest plot	
4 ^{1–4}	RCT	not serious	not serious	not serious	not serious	609/774 (78.7%)	268/384 (69.8%)	RR 1.12 (1.04 to 1.21)	84 more per 1000 (from 28 more to 147 more)	⊕⊕⊕⊕ HIGH
Any adverse e	vents –	donepezil (pro	bability of expe	riencing ≥1; f	ollow-up 10 to	24 week	s; lower is	s better)		
31,2,4	RCT	not serious	not serious	not serious	serious⁵	306/412 (74.3%)	141/205 (68.8%)	RR 1.07 (0.96 to 1.19)	48 more per 1000 (from 28 fewer to 131 more)	⊕⊕⊕O MODERATE
Any adverse e	vents -	rivastigmine (probability of ex	periencing ≥1	l; follow-up 24	weeks;	lower is b	etter)		
1 ³	RCT	not serious	N/A	not serious	not serious	303/362 (83.7%)	127/179 (70.9%)	RR 1.18 (1.06 to 1.31)	128 more per 1000 (from 43 more to 220 more)	⊕⊕⊕⊕ HIGH
Serious advers	se event	s – cholineste	erase inhibitors (probability of	experiencing	≥1; follo	w-up 24 v	veeks; lower is better);	see Figure 2 for forest plot	
2 ^{2,3}	RCT	not serious	serious ⁶	not serious	serious⁵	114/739 (15.4%)	48/352 (13.6%)	RR 1.13 (0.82 to 1.54)	18 more per 1000 (from 25 fewer to 74 more)	⊕⊕OO LOW
Serious advers	se event	s – donepezil	(probability of e	xperiencing ≥	1; follow-up 2	4 weeks;	lower is	better)		
1²	RCT	not serious	N/A	not serious	serious⁵	67/377 (17.8%)	22/173 (12.7%)	RR 1.4 (0.89 to 2.18)	51 more per 1000 (from 14 fewer to 150 more)	⊕⊕⊕O MODERATE
Serious advers	se event	s – rivastigmi	ne (probability o	of experiencing	g ≥1; follow-u	p 24 wee	ks; lower	is better)		
1 ³	RCT	not serious	N/A	not serious	serious⁵	47/362 (13%)	26/179 (14.5%)	RR 0.89 (0.57 to 1.39)	16 fewer per 1000 (from 62 fewer to 57 more)	⊕⊕⊕O MODERATE
Adverse event	s requiri	ing treatment	withdrawal - ch	olinesterase ir	nhibitors (prol	bability o	f experien	cing; follow-up 24 week	s; lower is better); see Figure 3 for forest plot	
3 ^{1–3}	RCT	not serious	not serious	not serious	not serious	122/753 (16.2%)	33/364 (9.1%)	RR 1.76 (1.23 to 2.53)	69 more per 1000 (from 21 more to 139 more)	⊕⊕⊕⊕ HIGH
Adverse event	s requiri	ing treatment	withdrawal - do	nepezil (proba	ability of exper	iencing;	follow-up	24 weeks)		
2 ^{1,2}	RCT	not serious	not serious	not serious	serious⁵	60/391 (15.3%)	19/185 (10.3%)	RR 1.46 (0.91 to 2.35)	47 more per 1000 (from 9 fewer to 139 more)	⊕⊕⊕O MODERATE
Adverse event	s requiri	ing treatment	withdrawal - riv	astigmine (pro	bability of ex	periencin	g; follow-	-up 24 weeks)		
1 ³	RCT	not serious	N/A	not serious	not serious	62/362 (17.1%)	14/179 (7.8%)	RR 2.19 (1.26 to 3.8)	93 more per 1000 (from 20 more to 219 more)	⊕⊕⊕⊕ HIGH
Hallucinations	– cholir	nesterase inhi	bitors (probabili	ty of experien	cing; follow-u	p 24 wee	ks; lower	is better); see Figure 4	for forest plot	
2 ^{2,3}	RCT	not serious	not serious	not serious	not serious	35/739 (4.7%)	31/352 (8.8%)	RR 0.54 (0.34 to 0.86)	41 fewer per 1000 (from 12 fewer to 58 fewer)	⊕⊕⊕⊕ HIGH
Hallucinations	– donep	oezil (probabil	ity of experienci	ng; follow-up	24 weeks; low	ver is bet	ter)			

Effoct

Error! No text of specified style in document.

		Quality	y assessment			No of p	patients		Effect	Quality	
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Chl	Placebo	Relative (95% CI)	Absolute (95% Cl)	Quanty	
1 ²	RCT	not serious	N/A	not serious	serious⁵	18/377 (4.8%)	14/173 (8.1%)	RR 0.59 (0.3 to 1.16)	33 fewer per 1000 (from 57 fewer to 13 more)	⊕⊕⊕O MODERATE	
Hallucinations	- rivast	igmine (proba	bility of experie	ncing; follow-u	up 24 weeks; I	ower is t	petter)				
1 ³	RCT	not serious	N/A	not serious	not serious	17/362 (4.7%)	17/179 (9.5%)	RR 0.49 (0.26 to 0.95)	48 fewer per 1000 (from 5 fewer to 70 fewer)	⊕⊕⊕⊕ HIGH	
1 Aarsland 2002 1 2 Dubois 2012; data for 2 active treatment groups were combined (donepezil 5mg and 10mg) 3 3 Emre 2004 4 4 Ravina 2005 5 5 At a 95% confidence level, data are consistent with appreciable harm, appreciable benefit or no difference											

PDD – rivastigmine patches vs. rivastigmine capsules: adverse events

		Qualit	y assessment			No of	patients	Effect			
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Rivastigmine patches	Rivastigmine capsules	Relative (95% Cl)	Absolute (95%Cl)	Quality	
Any adverse	events	(probability	of experiencin	ng ≥1; follow-	up 76 weeks	; lower is better)					
1 ¹	RCT	serious ²	N/A	not serious	not serious	263/288 (91.3%)	274/294 (93.2%)	RR 0.98 (0.93 to 1.03)	19 fewer per 1000 (from 65 fewer to 28 more)	⊕⊕OO LOW	
Serious adv	erse eve	nts (probab	ility of experie	ncing ≥1; fol	low-up 76 we	eks; lower is better)					
1 ¹	RCT	serious ²	N/A	not serious	serious ³	83/288 (28.8%)	87/294 (29.6%)	RR 0.97 (0.76 to 1.25)	9 fewer per 1000 (from 71 fewer to 74 more)	⊕⊕OO LOW	
Adverse eve	ents requ	uiring treatm	ent withdrawa	l (probability	of experienc	ing; follow-up 76 wee	ks; lower is better)				
1 ¹	RCT	serious ²	N/A	not serious	serious ³	71/288 (24.7%)	80/294 (27.2%)	RR 0.91 (0.69 to 1.19)	24 fewer per 1000 (from 84 fewer to 52 more)	⊕⊕OO LOW	
Hallucinatio	ns (prob	ability of ex	periencing ; fo	llow-up 76 w	eeks)						
1 ¹	RCT	serious ²	N/A	not serious	serious ³	25/288 (8.7%)	20/294 (6.8%)	RR 1.28 (0.73 to 2.25)	19 more per 1000 (from 18 fewer to 85 more)	⊕⊕OO LOW	

¹ Emre 2014

² Open-label study
 ³ Data are consistent with appreciable harm, appreciable benefit or no difference

	Cholinesterase in	nibitor	Place	bo		Risk Ratio	Risk Ratio
Study or Subgroup	Events	Total	Events	Total	Weight	M-H, Fixed, 95% Cl	M-H, Fixed, 95% Cl
1.1.1 Donepezil							
Aarsland 2002	12	14	9	12	2.7%	1.14 [0.77, 1.69]	
Dubois 2012	283	377	123	173	47.2%	1.06 [0.94, 1.18]	+
Ravina 2005 Subtotal (95% CI)	11	21 412	9	20 205	2.6% 52.5 %	1.16 [0.62, 2.19] 1.07 [0.96, 1.19]	•
Total events	306		141				
Heterogeneity: Chi ² =	0.22, df = 2 (P = 0.89	9); I ² = 09	6				
Test for overall effect	Z = 1.16 (P = 0.24)						
1.1.2 Rivastigmine							
Emre 2004 Subtotal (95% CI)	303	362	127	179 179	47.5% 47.5 %	1.18 [1.06, 1.31]	—
Total events	202	302	107	175	47.370	1.10[1.00, 1.51]	•
Heterogeneity: Not ar	onlicoblo		127				
Teet for overall effect:	7 – 211 /P – 0.002)						
restion overall ellect.	2 - 3.11 (1 - 0.002)						
Total (95% CI)		774		384	100.0%	1.12 [1.04, 1.21]	•
Total events	609 2.06 df = 2.78 = 0.50	2) - IZ — 100	268				
Tact for everall offect:	2.00, ui = 3 (F = 0.30	ŋ,ı = 05	0				'0.05 0.2 i Ś 20'
Test for subgroup dif	ferences: Chi ² = 1.78	, df = 1 (i	P = 0.18),	I ² = 43	.9%		Favours medication Favours placebo

PDD – cholinesterase inhibitor vs placebo: any adverse events (proportion of participants experiencing ≥1) – forest plot

	Cholinesterase inh	nibitor	Place	bo		Risk Ratio	Risk Ratio
Study or Subgroup	Events	Total	Events	Total	Weight	M-H, Fixed, 95% Cl	M-H, Fixed, 95% Cl
1.2.1 Donepezil							
Dubois 2012 Subtotal (95% CI)	67	377 377	22	173 173	46.4% 46. 4%	1.40 [0.89, 2.18] 1.40 [0.89, 2.18]	
Total events	67		22				
Heterogeneity: Not app	plicable						
Test for overall effect: 2	Z = 1.47 (P = 0.14)						
1.2.2 Rivastigmine							
Emre 2004 Subtotal (95% CI)	47	362 362	26	179 179	53.6% 53.6 %	0.89 [0.57, 1.39] 0.89 [0.57, 1.39]	
Total events	47		26				
Heterogeneity: Not app	plicable						
Test for overall effect: 2	Z = 0.49 (P = 0.62)						
Total (95% CI)		739		352	100.0%	1.13 [0.82, 1.54]	•
Total events	114		48				
Heterogeneity: Chi ² = 1	1.94, df = 1 (P = 0.16	6); I ² = 48	1%				
Test for overall effect: 2	Z = 0.75 (P = 0.45)						Eavours medication Eavours placebo
Test for subgroup diffe	erences: Chi ^z = 1.93	. df = 1 (l	P = 0.16).	l² = 48	.3%		ravous medication ravous placebo

PDD – cholinesterase inhibitor vs placebo: serious adverse events (proportion of participants experiencing ≥1) – forest plot

	Cholinesterase inh	ibitor	Place	bo		Risk Ratio	Risk Ratio
Study or Subgroup	Events	Total	Events	Total	Weight	M-H, Fixed, 95% Cl	M-H, Fixed, 95% Cl
1.3.1 Donepezil							
Aarsland 2002	2	14	0	12	1.2%	4.33 [0.23, 82.31]	
Dubois 2012 Subtotal (95% CI)	58	377 391	19	173 185	57.5% 58.7 %	1.40 [0.86, 2.28] 1.46 [0.91, 2.35]	
Total events	60		19			. , ,	-
Heterogeneity: Chi ² =	0.55, df = 1 (P = 0.46)); I ² = 09	6				
Test for overall effect:	Z = 1.55 (P = 0.12)	•					
1.3.2 Rivastigmine							
Emre 2004 Subtotal (95% CI)	62	362 362	14	179 179	41.3% 41.3 %	2.19 [1.26, 3.80] 2.19 [1.26, 3.80]	
Total events	62		14				
Test for overall effect:	Z = 2.79 (P = 0.005)						
Total (95% CI)		753		364	100.0%	1.76 [1.23, 2.53]	◆
Total events	122		33				
Heterogeneity: Chi ² =	1.81, df = 2 (P = 0.40)); I ≃ = 09	6				
Test for overall effect:	Z = 3.08 (P = 0.002)						U.05 U.2 1 5 20
Test for subgroup diff	ferences: Chi ² = 1.19,	df = 1 (i	P = 0.28),	, I² = 15	.7%		ravours medication ravours placebo

PDD – cholinesterase inhibitor vs placebo: adverse events requiring treatment withdrawal (proportion of participants experiencing) forest plot

	Cholinesterase inhil	bitor	Place	bo		Risk Ratio	Risk Ratio
Study or Subgroup	Events	Total	Events	Total	Weight	M-H, Fixed, 95% Cl	M-H, Fixed, 95% Cl
1.4.1 Donepezil							
Dubois 2012	18	377	14	173	45.8%	0.59 [0.30, 1.16]	
Subtotal (95% CI)		377		173	45.8%	0.59 [0.30, 1.16]	
Total events	18		14				
Heterogeneity: Not ap	oplicable						
Test for overall effect:	: Z = 1.53 (P = 0.13)						
1.4.2 Rivastigmine							
Emre 2004	17	362	17	179	54.2%	0.49 [0.26, 0.95]	
Subtotal (95% CI)		362		179	54.2%	0.49 [0.26, 0.95]	
Total events	17		17				
Heterogeneity: Not ap	oplicable						
Test for overall effect:	: Z = 2.13 (P = 0.03)						
T-4-1/05% ON		700		959	400.00		
l otal (95% CI)		739		352	100.0%	0.54 [0.34, 0.86]	-
Total events	35		31				
Heterogeneity: Chi ² =	: 0.14, df = 1 (P = 0.71);	² = 09	6				
Test for overall effect:	: Z = 2.60 (P = 0.009)						Favours medication Favours placebo
Test for subgroup dif	ferences: Chi² = 0.14, (df = 1 (i	P = 0.71),	l² = 0%	6		rate and the stear of the other place to be

PDD – cholinesterase inhibitor vs placebo: hallucinations (proportion of participants experiencing) – forest plot

PDD – cholinesterase inhibitor vs. placebo: cognitive function

		Qual	ity assessment			No of	patients	Effect	Quality
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Chl	Placebo	Mean difference (95% Cl)	Quanty
MMSE – cholinest	erase inhi	bitors (follow-up	10 to 24 weeks; ran	ige of scores: 0-3	0; higher is better)	; see Fig	gure 5 for	forest plot	
41-4	RCT	not serious	not serious	not serious	not serious	752	367	1.36 higher (0.95 to 1.77 higher)	⊕⊕⊕⊕ HIGH
MMSE – donepezi	l (follow-u	p 10 to 24 weeks	; range of scores: 0	-30; higher is bett	ter)				
31,2,4	RCT	not serious	not serious	not serious	not serious	417	201	1.58 higher (1.06 to 2.1 higher)	⊕⊕⊕⊕ HIGH
MMSE – rivastigm	ine (follow	v-up 24 weeks; ra	ange of scores: 0-30	; higher is better)	1				
1 ³	RCT	not serious	N/A	not serious	not serious	335	166	1 higher (0.33 to 1.67 higher)	⊕⊕⊕⊕ HIGH
ADAS-cog – choli	nesterase	inhibitors (follow	v-up 10 to 24 weeks	; range of scores:	0-70; lower is bett	er); se	e Figure 6	for forest plot	
3 ^{1,2,4}	RCT	not serious	not serious	not serious	not serious	689	346	2.28 lower (3.40 to 1.15 lower)	⊕⊕⊕⊕ HIGH
ADAS-cog – done	pezil (follo	ow-up 10 to 24 we	eeks; range of score	es: 0-70; lower is l	better)				
2 ^{2,4}	RCT	not serious	not serious	not serious	serious⁵	360	185	1.5 lower (3.28 lower to 0.27 higher)	⊕⊕⊕O MODERATE
ADAS-cog – rivas	tigmine (fo	ollow-up 24 week	s; range of scores:	0-70; lower is bet	ter)				
1 ³	RCT	not serious	N/A	not serious	not serious	329	161	2.8 lower (4.26 to 1.34 lower)	⊕⊕⊕⊕ HIGH
MDRS (total score) – choline	esterase inhibito	rs (follow-up 10 to 2	4 weeks; range o	f scores: 0-144; hig	gher is b	oetter) ⁶ se	e Figure 7 for forest plot	
2 ^{3,4}	RCT	not serious	not serious	not serious	very serious5,7	35	31	3.39 higher (4.06 lower to 10.84 higher)	⊕⊕OO LOW
MDRS (total score) – donep	ezil (follow-up 10	weeks; range of sc	ores: 0-144; high	er is better)				
14	RCT	not serious	N/A	not serious	very serious5,7	19	19	0.2 lower (11.44 lower to 11.04 higher)	⊕⊕OO LOW
MDRS (total score) – rivasti	gmine (follow-up	24 weeks; range of	scores: 0-144; hi	gher is better)6				
1 ³	RCT	serious ⁷	N/A	not serious	serious⁵	16	12	6.21 higher (3.75 lower to 16.17 higher)	⊕⊕OO LOW
Clock drawing tes	t – rivastig	gmine (follow-up	24 weeks; range of	scores: 0-10; hig	her is better)				
1 ³	RCT	serious ⁷	N/A	not serious	serious⁵	49	30	1.1 higher (0.01 lower to 2.21 higher)	⊕⊕OO LOW
D-KEFS verbal flu	ency test	(total score) – riv	vastigmine (follow-u	p 24 weeks; meas	ured by number of	f correc	t respons	es; higher is better)	
1 ³	RCT	not serious	N/A	not serious	not serious	258	144	2.8 higher (1.47 to 4.13 higher)	⊕⊕⊕⊕ HIGH
D-KEFS verbal flu	ency test	(letter fluency) -	donepezil (follow-u	o 24 weeks; highe	er is better)				
1 ²	RCT	not serious	N/A	not serious	not serious	307	152	2.83 higher (0.95 to 4.71 higher)	⊕⊕⊕⊕ HIGH
D-KEFS verbal flu	ency test	(category fluency	/) – donepezil (follow	w-up 24 weeks; hi	igher is better)				

	Quali	ty assessment			No of	patients	Effect	Quality
Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Chl	Placebo	Mean difference (95% CI)	Quanty
RCT	not serious	N/A	not serious	not serious	307	152	3.93 higher (2.05 to 5.81 higher)	⊕⊕⊕⊕ HIGH
ency test (category switchi	ng) – donepezil (fol	low-up 24 weeks;	higher is better)				
RCT	not serious	N/A	not serious	serious⁵	307	152	1.09 higher (0.79 lower to 2.97 higher)	⊕⊕⊕O MODERATE
e (follow-u	ıp 24 weeks; mea	sured with: millised	conds; lower is be	etter)				
RCT	not serious	N/A	not serious	serious⁵	328	158	173.7 lower (471.23 lower to 123.83 higher)	⊕⊕⊕O MODERATE
ollow-up 2	24 weeks; range	of scores: 0-20; hig	her is better)					
RCT	serious ⁸	N/A	not serious	not serious	221	111	0.88 higher (0.4 to 1.37 higher)	⊕⊕⊕O MODERATE
	Design RCT ency test (RCT e (follow-u RCT ollow-up 2 RCT	Quality Design Risk of bias RCT not serious ency test (category switching) RCT not serious e (follow-up 24 weeks; means RCT not serious ollow-up 24 weeks; range of RCT RCT serious ⁸	Quality assessmentDesignRisk of biasInconsistencyRCTnot seriousN/Aency test (category switching) – donepezil (folRCTnot seriousN/Ae (follow-up 24 weeks; measured with: millisedRCTnot seriousN/Aollow-up 24 weeks; rangeof scores: 0-20; highRCTserious ⁸ N/A	Quality assessment Design Risk of bias Inconsistency Indirectness RCT not serious N/A not serious ency test (category switching) – donepezil (follow-up 24 weeks; not serious RCT RCT not serious N/A not serious e (follow-up 24 weeks; measured with: milliseconds; lower is bet RCT not serious N/A not serious N/A not serious N/A ollow-up 24 weeks; range of scores: 0-20; higher is better) RCT serious ⁸	Quality assessmentDesignRisk of biasInconsistencyIndirectnessImprecisionRCTnot seriousN/Anot seriousnot seriousency test (category switching) – donepezil (follow-up 24 weeks; higher is better)not seriousserious ⁵ RCTnot seriousN/Anot seriousserious ⁵ et (follow-up 24 weeks; measured with: milliseconds; lower is better)not seriousserious ⁵ RCTnot seriousN/Anot seriousserious ⁵ ollow-up 24 weeks; rangeof scores: 0-20; higher is better)not seriousnot seriousRCTserious ⁸ N/Anot seriousnot serious	Quality assessmentNo ofDesignRisk of biasInconsistencyIndirectnessImprecisionChlRCTnot seriousN/Anot seriousnot serious307ency test (category switching) – donepezil (follow-up 24 weeks; higher is better)RCTnot seriousN/Anot seriousserious ⁵ 307RCTnot seriousN/Anot seriousserious ⁵ 307e (follow-up 24 weeks; measured with: milliseconds; lower is better)RCTnot seriousserious ⁵ 328ollow-up 24 weeks; rangeof scores: 0-20; higher is better)serious ⁵ 328ollow-up 24 weeks; rangeN/Anot seriousnot serious221	Quality assessmentNo of patientsDesignRisk of biasInconsistencyIndirectnessImprecisionChlPlaceboRCTnot seriousN/Anot seriousnot serious307152ency test (category switching) – donepezil (follow-up 24 weeks; higher is better)RCTnot seriousN/Anot seriousserious ⁵ 307152e (follow-up 24 weeks; measured with: milliseconds; lower is better) </td <td>Quality assessmentNo of patientsEffectDesignRisk of biasInconsistencyIndirectnessImprecisionChiPlaceboMean difference (95% Cl)RCTnot seriousN/Anot seriousnot serious3071523.93 higher (2.05 to 5.81 higher)ency test (category switch:rug) – donepezil (follow-up 24 weeks; higher is better)not serious3071521.09 higher (0.79 lower to 2.97 higher)RCTnot seriousN/Anot seriousserious⁵3071521.09 higher (0.79 lower to 2.97 higher)e (follow-up 24 weeks; mesured with: milliseconds; lower is better)serious⁵328158173.7 lower (471.23 lower to 123.83 higher)RCTnot seriousN/Anot seriousserious⁵328158173.7 lower (471.23 lower to 123.83 higher)ollow-up 24 weeks; ranger scores: 0-20; higher is better)serious2211110.88 higher (0.4 to 1.37 higher)</td>	Quality assessmentNo of patientsEffectDesignRisk of biasInconsistencyIndirectnessImprecisionChiPlaceboMean difference (95% Cl)RCTnot seriousN/Anot seriousnot serious3071523.93 higher (2.05 to 5.81 higher)ency test (category switch:rug) – donepezil (follow-up 24 weeks; higher is better)not serious3071521.09 higher (0.79 lower to 2.97 higher)RCTnot seriousN/Anot seriousserious ⁵ 3071521.09 higher (0.79 lower to 2.97 higher)e (follow-up 24 weeks; mesured with: milliseconds; lower is better)serious ⁵ 328158173.7 lower (471.23 lower to 123.83 higher)RCTnot seriousN/Anot seriousserious ⁵ 328158173.7 lower (471.23 lower to 123.83 higher)ollow-up 24 weeks; ranger scores: 0-20; higher is better)serious2211110.88 higher (0.4 to 1.37 higher)

¹ Aarsland 2002

² Dubois 2012; data for 2 active treatment groups were combined (donepezil 5mg and 10mg). Mean and standard deviation calculated from data reported in paper

³ Emre 2004

⁴ Ravina 2005

⁵ At a 95% confidence level, data are consistent with appreciable harm, appreciable benefit or no difference

⁶ Data from Emre 2004 reported in a secondary publication (Dujardin 2006)
 ⁷ Small numbers of participants in the analysis

⁸ Data available for only a small proportion of all participants for this outcome

PDD - rivastigmine patches vs. rivastigmine capsules: cognitive outcomes

	- J				<u> </u>			J					
		Qualit	y asses	sment					No of	patients		Effect	Quality
No of studies	Design	Risk of bias	Incons	sistency	y Indir	ectness	s Imp	recision	Rivastigmine patches	Rivastigmine caps	sules	Mean difference (95% CI)	Quanty
MDRS (total sc	ore) (foll	ow-up 24 we	eks; rang	ge of so	cores 0-	144; hig	gher is	better)					
1 ¹	RCT	serious ²	N/A		not se	erious	serio	US ³	273	273		2.1 lower (4.27 lower to 0.07 higher) ⊕⊕OO LOW
MDRS (total sc	ore) (foll	ow-up 76 we	eks; rang	ge of so	cores 0-	144; hig	gher is	better)					
1 ¹	RCT	serious ²	N/A		not se	erious	not s	erious	273	273		5.3 lower (8.17 to 2.43 lower)	⊕⊕⊕O MODERATE
¹ Emre 2014 ² Open-label s ³ At a 95% coi	tudy nfidence	level, data a	are cons	sistent	with app	preciab	le har	m, appre	eciable benefit or no dit	ference			
		Cholinestera	se inhibi	tor	Pla	acebo			Mean Difference	Mean Diff	erence	e	
Study or Subg	roup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Fixed, 95% Cl	IV, Fixed,	95% C	<u> </u>	
1.5.1 Donepez	il												
Aarsland 2002		22.8	3.7	12	21	5	12	1.4%	1.80 [-1.72, 5.32]				
Dubois 2012		22.974 3	.009	386	21.427	3.009	170	57.4%	1.55 [1.00, 2.09]				
Ravina 2005 Subtotal (95%	CI)	24.5	3.2	19 417	22.5	3.7	19 201	3.5% 62.3 %	2.00 [-0.20, 4.20] 1.58 [1.06, 2.10]	Ť	٠		
Heterogeneity: Test for overall	Chi² = 0. I effect: Z	17, df = 2 (P = = 5.93 (P < 0.)	: 0.92); I² 00001)	= 0%									
1.5.2 Rivastigr	nine												
Emre 2004 Subtotal (95% Heterogeneity: Test for overall	CI) Not appl I effect: Z	0.8 licable = 2.92 (P = 0.1	3.8 003)	335 335	-0.2	3.5	166 166	37.7% 37.7 %	1.00 (0.33, 1.67) 1.00 (0.33, 1.67)		•		
Total (95% CI)				752			367	100.0%	1.36 [0.95, 1.77]		٠		

-10

Test for overall effect: Z = 6.48 (P < 0.00001)

Heterogeneity: Chi² = 1.95, df = 3 (P = 0.58); l² = 0%

Test for subgroup differences: Chi² = 1.78, df = 1 (P = 0.18), l² = 43.8%

PDD – cholinesterase inhibitor vs placebo: MMSE – forest plot

-5

ń.

Favours placebo Favours medication

10

<u>5</u>

	Cholines	sterase inhi	ibitor	P	lacebo			Mean Difference	Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Fixed, 95% Cl	IV, Fixed, 95% Cl
1.6.1 Donepezil									
Dubois 2012	19.77	10.165	341	21.22	10.165	166	35.7%	-1.45 [-3.34, 0.44]	
Ravina 2005 Subtotal (95% CI)	22.5	6.9	19 360	24.4	9.4	19 185	4.6% 40.3 %	-1.90 [-7.14, 3.34] - 1.50 [-3.28, 0.27]	
Heterogeneity: Chi² = Test for overall effect:	0.03, df = 1 Z = 1.66 (P	(P = 0.87); = 0.10)	I² = 0%						
1.6.2 Rivastigmine									
Emre 2004 Subtotal (95% CI)	-2.1	8.2	329 329	0.7	7.5	161 161	59.7% 59.7 %	-2.80 [-4.26, -1.34] - 2.80 [-4.26, -1.34]	
Heterogeneity: Not ap Test for overall effect:	oplicable Z = 3.76 (P	= 0.0002)							
Total (95% CI)			689			346	100.0%	-2.28 [-3.40, -1.15]	•
Heterogeneity: Chi ² =	1.25, df = 2	? (P = 0.53);	l²=0%						
Test for overall effect:	Z = 3.96 (P	< 0.0001)							-10 -5 0 5 10 Envours modication Envours placabo
Test for subgroup diff	, ferences: C	hi² = 1.23, o	#f = 1 (P :	= 0.27),	l [≈] = 18.69	%			ravours medication ravours placebo

PDD – cholinesterase inhibitor vs placebo: ADAS-cog – forest plot

	Cholinest	terase inhi	ibitor	Р	lacebo			Mean Difference	Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Fixed, 95% Cl	IV, Fixed, 95% Cl
1.7.1 Donepezil									
Ravina 2005	108.3	17.13	19	108.5	18.2	19	44.0%	-0.20 [-11.44, 11.04]	_
Subtotal (95% CI)			19			19	44.0%	-0.20 [-11.44, 11.04]	
Heterogeneity: Not ap	plicable								
Test for overall effect:	Z = 0.03 (P	= 0.97)							
1.7.2 Rivastigmine									
Emre 2004	5.79	12.99	16	-0.42	13.54	12	56.0%	6.21 [-3.75, 16.17]	
Subtotal (95% Cl)			16			12	56.0%	6.21 [-3.75, 16.17]	
Heterogeneity: Not ap	plicable								
Test for overall effect:	Z=1.22 (P	= 0.22)							
Total (95% CI)			35			31	100.0%	3.39 [-4.06, 10.84]	
Heterogeneity: Chi ² =	0.70, df = 1	(P = 0.40);	I² = 0%					-	
Test for overall effect:	Z = 0.89 (P :	= 0.37)							-20 -10 0 10 20 Eavours placebol Eavours modication
Test for subgroup diff	erences: Ch	ni² = 0.70, c	lf = 1 (P =	= 0.40),	l² = 0%				ravouis placebo ravouis illeuication

PDD – cholinesterase inhibitor vs placebo: MDRS (total score) – forest plot

	10010100	Quality	assessment		one	No of natie	onts		
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Chl	Placebo	Effect (95%CI)	Quality
Global function –	cholineste	erase inhibitors (follow-up 10 to 24 w	eeks: measured w	vith: CIBIC+. A	DCS-CGIC or CO	GIC: range	of scores: 1-7: lower is better): see Figure 8 f	or forest plot
4 ^{1–4}	RCT	not serious	not serious	not serious	not serious	707	366	SMD 0.3 lower (0.42 to 0.17 lower)	⊕⊕⊕⊕ HIGH
Global response plot	- cholinest	terase inhibitors	(at least minimal im	provement; follow	-up 10 to 24 w	eeks; measured	with: CIB	IC+ or ADCS-CGIC; higher is better); see Figu	re 9 for forest
3 ^{1–3}	RCT	not serious	not serious	not serious	not serious	294/688 (42.7%)	119/347 (34.3%)	RR 1.24 (1.05 to 1.47) 82 more per 1000 (from 17 more to 161 more)	⊕⊕⊕⊕ HIGH
Global response	– donepezi	il (at least minima	al improvement; folle	ow-up 10 to 24 we	eks; measured	d with: CIBIC+; h	igher is b	etter)	
2 ^{1,2}	RCT	not serious	not serious	not serious	serious⁵	160/359 (44.6%)	70/182 (38.5%)	RR 1.15 (0.92 to 1.42) 58 more per 1000 (from 31 fewer to 162 more)	⊕⊕⊕O MODERATE
Global response	- rivastigm	nine (at least mini	mal improvement; f	ollow-up 24 week	s; measured w	ith: ADCS-CGIC	; higher is	s better)	
1 ³	RCT	not serious	N/A	not serious	not serious	134/329 (40.7%)	49/165 (29.7%)	RR 1.37 (1.05 to 1.79) 110 more per 1000 (from 15 more to 235 more)	⊕⊕⊕⊕ HIGH
CIBIC+ – donepe	zil (follow-	up 10 to 24 week	s; range of scores:	1-7; lower is bette	r); see Figure [,]	10 for forest plot			
21,2	RCT	not serious	serious ⁶	not serious	serious⁵	359	182	MD 0.43 lower (0.93 lower to 0.08 higher)	⊕⊕OO LOW
CGIC – donepezi	(follow-up	10 weeks; range	of scores: 1-7; low	er is better)					
14	RCT	not serious	N/A	not serious	very serious ^{5,7}	19	19	MD 0.37 lower (0.89 lower to 0.15 higher)	⊕⊕OO LOW
UPDRS (total sco	re) – done	pezil (follow-up 1	0 weeks; range of s	cores: 0-199; lowe	er is better)				
14	RCT	not serious	N/A	not serious	very serious ^{5,7,8}	21	20	MD 2.3 lower (15.77 lower to 11.17 higher)	⊕⊕OO LOW
ADCS-CGIC - riv	astigmine (follow-up 24 wee	ks; range of scores	: 1-7; lower is bet	ter)				
1 ³	RCT	not serious	N/A	not serious	not serious	329	165	MD 0.5 lower (0.77 to 0.23 lower)	⊕⊕⊕⊕ HIGH
¹ Aarsland 2002 ² Dubaia 2012: c	lata far 2 c	ativo trootmont	around word comb	inad (dananazil E	ima and 10ma	x) Moon and at	ndord d	aviation coloulated from data reported in par	or

abolinactoreas inhibitor va placeba; glabal accessment חחם

² Dubois 2012; data for 2 active treatment groups were combined (donepezil 5mg and 10mg). Mean and standard deviation calculated from data reported in paper ³ Emre 2004

⁴ Ravina 2005

⁵ At a 95% confidence level, data are consistent with appreciable harm, appreciable benefit or no difference

 $^{6}i^{2} > 40\%$ between studies

⁷ Data from a single very small study
 ⁸CI cross MID of 7.3 points (Schrag et al., 2006)

	Cholinesterase inhibitor		P	acebo		5	Std. Mean Difference	Std. Mean Difference		
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Fixed, 95% Cl	IV, Fixed, 95% CI	
1.15.1 Donepezil										
Aarsland 2002	3.3	0.9	12	4.1	0.8	12	2.3%	-0.91 [-1.76, -0.06]		
Dubois 2012	3.65	1.2	347	3.9	1.27	170	48.0%	-0.20 [-0.39, -0.02]	•	
Ravina 2005 Subtotal (95% CI)	3.58	0.77	19 378	3.95	0.85	19 201	3.9% 54.2 %	-0.45 [-1.09, 0.20] - 0.25 [-0.42, -0.08]		
Heterogeneity: Chi ² =	2.90, df = 2	(P = 0.23);	I ² = 31%	,						
Test for overall effect:	Z = 2.84 (P =	= 0.005)								
1.15.2 Rivastigmine										
Emre 2004 Subtotal (95% CI)	3.8	1.4	329 329	4.3	1.5	165 165	45.8% 4 5.8 %	-0.35 [-0.54, -0.16] - 0.35 [-0.54, -0.16]		
Heterogeneity: Not ap	plicable									
Test for overall effect:	Z = 3.62 (P =	= 0.0003)								
Total (95% CI)			707			366	100.0%	-0.30 [-0.42, -0.17]	•	
Heterogeneity: Chi ² =	3.46, df = 3	(P = 0.33);	I² = 13%	,						
Test for overall effect:	Z = 4.54 (P <	< 0.00001))						-4 -2 U 2 4 Eaveure modication Eaveure placebo	
Test for subaroup diff	erences: Ch	i² = 0.56. c	; :f=1 (P=	= 0.46).	l ^z = 0%	6			Favours medication Favours placebo	

PDD – cholinesterase inhibitor vs placebo: global function (different measures)

	Cholinesterase in	hibitor	Place	bo		Risk Ratio	Risk Ratio
Study or Subgroup	Events	Total	Events	Total	Weight	M-H, Fixed, 95% Cl	M-H, Fixed, 95% Cl
1.16.1 Donepezil							
Aarsland 2002	5	12	2	12	1.3%	2.50 [0.60, 10.46]	
Dubois 2012	155	347	68	170	57.6%	1.12 [0.90, 1.39]	-
Subtotal (95% CI)		359		182	58.8 %	1.15 [0.92, 1.42]	◆
Total events	160		70				
Heterogeneity: Chi ² = 1	1.20, df = 1 (P = 0.2)	7); I ² = 16	6%				
Test for overall effect: 2	Z = 1.24 (P = 0.21)						
1.16.2 Rivastigmine							
Emre 2004	134	329	49	165	41.2%	1.37 [1.05, 1.79]	
Subtotal (95% CI)		329		165	41.2 %	1.37 [1.05, 1.79]	◆
Total events	134		49				
Heterogeneity: Not app	olicable						
Test for overall effect: 2	Z = 2.31 (P = 0.02)						
Total (95% CI)		688		347	100.0%	1.24 [1.05, 1.47]	◆
Total events	294		119				
Heterogeneity: Chi ² = 2	2.35, df = 2 (P = 0.3	1); I ^z = 16	5%				
Test for overall effect: 2	Z = 2.49 (P = 0.01)						0.05 0.2 1 5 20 Eavours placebol Eavours modication
Test for subgroup diffe	erences: Chi ² = 1.04	l, df = 1 (l	P = 0.31).	$ ^{2} = 4.0$)%		Favours placebo Favours medication

PDD – cholinesterase inhibitor vs placebo: global response (at least minimal improvement) – forest plot

	Cholineste	Pla	acebo			Mean Difference	Mean Difference		
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Random, 95% Cl	IV, Random, 95% Cl
1.17.1 Donepezil									
Aarsland 2002	3.3	0.9	12	4.1	0.8	12	32.3%	-0.80 [-1.48, -0.12]	
Dubois 2012 Subtotal (95% Cl)	3.65	1.2	347 359	3.9	1.27	170 182	67.7% 100.0 %	-0.25 [-0.48, -0.02] - 0.43 [-0.93, 0.08]	•
Heterogeneity: Tau ² = Test for overall effect:	= 0.08; Chi² = : : Z = 1.66 (P =	2.25, df = 0.10)	1 (P = 0.	13); I ^z =	56%				
Total (95% CI)			359			182	100.0 %	-0.43 [-0.93, 0.08]	•
Heterogeneity: Tau² = Test for overall effect: Test for subgroup diff	= 0.08; Chi² = : : Z = 1.66 (P = ferences: Not	2.25, df = 0.10) applicabl	1 (P = 0. e	13); I² =	56%			-	-4 -2 0 2 4 Favours medication Favours placebo

PDD – cholinesterase inhibitor (donepezil) vs placebo: CIBIC+ – forest plot

[Insert footer here]

PDD – cholinesterase inhibitor vs. placebo: activities of daily living

		Quality	assessment			No of p	oatients		Quality				
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Chl	Placebo		Quanty				
ADL – cholinesterase inhibitors (follow-up 24 weeks; measured with: ADCS-ADL or DAD; higher is better); see Figure 11 for forest plot													
2 ^{1,2}	RCT	not serious	not serious	not serious	not serious	684	335	SMD 0.18 higher (0.05 to 0.31 higher)	⊕⊕⊕⊕ HIGH				
DAD – donepezil (follow-up 24 weeks; range of scores 0-100; higher is better)													
1 ¹	RCT	not serious	N/A	not serious	serious ³	351	170	MD 2.26 higher (0.38 lower to 4.89 higher)	⊕⊕⊕O MODERATE				
ADCS-ADL – rivastig	mine (follow	/-up 24 weeks; rang	ge of scores: 0-78; high	er is better)									
1²	RCT	not serious	N/A	not serious	not serious	333	165	MD 2.5 higher (0.43 to 4.57 higher)	⊕⊕⊕⊕ HIGH				
¹ Dubois 2012; data ² Emre 2004	for 2 active	e treatment groups	were combined (don	epezil 5mg and 10r	ng). Mean and	d standard	deviation	calculated from data reported in paper					

³ At a 95% confidence level, data are consistent with appreciable harm, appreciable benefit or no difference

PDD - rivastigmine patches vs. rivastigmine capsules: activities of daily living

		Quality	y assessment			No of	patients	Effect	Quality				
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Rivastigmine patches	Rivastigmine capsules	Mean difference (95% Cl)	Quanty				
ADCS-ADL (follow-up 24 weeks; range of scores: 0-78; higher is better)													
1 ¹	RCT	serious ²	N/A	not serious	serious ³	270	273	0.9 lower (2.67 lower to 0.87 higher)	⊕⊕OO LOW				
ADCS-ADL (fo	llow-up	76 weeks; rang	ge of scores: 0-7	78; higher is b	etter)								
1 ¹	RCT	serious ²	N/A	not serious	not serious	270	273	3.4 lower (5.84 to 0.96 lower)	⊕⊕⊕O MODERATE				
¹ Emre 2014 ² Open-label ³ At a 95% co	Emre 2014 Open-label study At a 95% confidence level, data are consistent with appreciable harm, appreciable benefit or no difference												

	Favo	urs place	ebo	F	Placebo			Std. Mean Difference	Std. Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Fixed, 95% Cl	IV, Fixed, 95% Cl
1.21.1 Donepezil									
Dubois 2012 Subtotal (95% CI)	0	14.386	351 351	-2.26	14.386	170 170	51.0% 51.0 %	0.16 [-0.03, 0.34] 0.16 [-0.03, 0.34]	•
Heterogeneity: Not a	pplicable	l							
Test for overall effect	:Z=1.68) (P = 0.09	3)						
1.21.2 Rivastigmine									
Emre 2004 Subtotal (95% CI)	-1.1	12.6	333 333	-3.6	10.3	165 165	49.0% 49.0 %	0.21 [0.02, 0.40] 0.21 [0.02, 0.40]	 ◆
Heterogeneity: Not a	pplicable	l							
Test for overall effect	: Z = 2.20) (P = 0.03	3)						
Total (95% CI)			684			335	100.0%	0.18 [0.05, 0.31]	◆
Heterogeneity: Chi ² =	: 0.16, df	= 1 (P = 0).69); I ²	= 0%					
Test for overall effect	: Z = 2.74	(P = 0.0)	J6) Ü						-2 -1 U 1 2
Test for subgroup dif	ferences	Chi² = 0	16, df:	= 1 (P =	0.69), I ^z :	= 0%			Favours placebo Favours medication

PDD – cholinesterase inhibitor vs placebo: ADL (different measures) – forest plot

PDD – cholinesterase	inhibitor vs.	placebo: other non	-cognitive outcomes

		Quality	assessment			No of patients		Effect	Quality			
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Chl	Placebo	Mean difference (95% CI)	Quanty			
NPI-10 item – choli	nesterase i	nhibitors (follow-u	up 24 weeks; range of	f scores: 0-120; lov	wer is better);	see Fig	gure 12 for	forest plot				
2 ^{1,2}	RCT	not serious ³	not serious	not serious	not serious	688	336	1.67 lower (3.01 to 0.32 lower)	⊕⊕⊕⊕ HIGH			
PI-10 item – donepezil (follow-up 24 weeks; range of scores: 0-120; lower is better)												
1 ¹	RCT	not serious ³	N/A	not serious	serious ⁴	354	170	1.34 lower (3.23 lower to 0.54 higher)	⊕⊕⊕O MODERATE			
NPI-10 item – rivas	tigmine (fo	llow-up 24 weeks;	range of scores: 0-12	20; lower is better)								
1 ²	RCT	not serious	N/A	not serious	not serious	334	166	2.00 lower (3.91 to 0.09 lower)	⊕⊕⊕⊕ HIGH			
UPDRS III – donepe	zil (follow-	up 10 weeks; low	er is better); see Figu	re 13 for forest plo	t							
25,6	RCT	serious ⁷	not serious	not serious	serious ^{4,8}	33	32	1.5 lower (7.87 lower to 4.87 higher)	⊕⊕OO LOW			
¹ Dubois 2012; da ² Emre 2004 ³ Data for this out downgraded ⁴ At a 05% confide	ta for 2 act come not r	tive treatment gro eported in Aarsla	oups were combined nd 2002. This repres	(donepezil 5mg a sents a very smal	and 10mg). N I proportion o	fean ar of the to	nd standa otal partici	rd deviation calculated from data reported in pa pants in the analysis, therefore quality assessm	per ent not			

⁴ At a 95% confic ⁵ Aarsland 2002

6 Ravina 2005

⁷Data for this outcome not reported in 2 large RCTs (Dubois 2012 and Emre 2004). Papers stated no significant difference between groups ⁸CI cross MID between 3.25 (Horvath et al., 2015) and 5 points (Schrag et al., 2006)

PDD – rivastigmine patches vs. rivastigmine capsules: other non-cognitive outcomes

						U							
		Quality	/ assessment			No of	patients	Effect	Quality				
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Rivastigmine patches	Rivastigmine capsules	Mean difference (95% Cl)	Quanty				
NPI-10 item (follow-up 24 weeks; range of scores: 0-120; lower is better)													
1 ¹	RCT	serious ²	N/A	not serious	serious ³	273	273	1.6 higher (0.13 lower to 3.33 higher)	⊕⊕OO LOW				
NPI-10 item (follow-up 76 weeks; range of scores: 0-120; lower is better)													
1 ¹	RCT	serious ²	N/A	not serious	not serious	273	273	2.3 lower (4.3 to 0.3 lower)	⊕⊕⊕O MODERATE				
UPDRS III (foll	ow-up 76	weeks; lower	is better)										
1 ¹	RCT	serious ²	N/A	not serious	not serious ⁴	175	183	0 higher (2.04 lower to 2.04 higher)	⊕⊕⊕O MODERATE				
1													

¹ Emre 2014

² Open-label study

³ At a 95% confidence level, data are consistent with appreciable harm, appreciable benefit or no difference

⁴CI do not cross MID between 3.25 (Horvath et al., 2015) and 5 points (Schrag et al., 2006)

	Cholinest	ibitor	Placebo				Mean Difference	Mean Difference	
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Fixed, 95% Cl	I IV, Fixed, 95% CI
1.26.1 Donepezil									
Dubois 2012 Subtotal (95% CI)	11.711	10.32	354 354	13.055	10.32	170 170	50.6% 50.6 %	-1.34 [-3.23, 0.54] - 1.34 [-3.23, 0.54]	
Heterogeneity: Not app	olicable								
Test for overall effect: 2	Z = 1.40 (P :	= 0.16)							
1.26.2 Rivastigmine									
Emre 2004 Subtotal (95% Cl)	-2	10	334 334	0	10.4	166 166	49.4% 49. 4%	-2.00 [-3.91, -0.09] - 2.00 [-3.91, -0.09]	
Heterogeneity: Not app	olicable								
Test for overall effect: 2	Z = 2.05 (P :	= 0.04)							
Total (95% CI)			688			336	100.0%	-1.67 [-3.010.32]	•
Heterogeneity: Chi ² = (Test for overall effect: 2 Test for subgroup diffe	² = 0% f = 1 (P -	- 0 6 3) IZ	- 0%				-20 -10 0 10 20 Favours medication Favours placebo		
reaction candidap ante		n = 0.20, 0	a = 10	0.007,1	- 0 /0				

PDD – cholinesterase inhibitor vs placebo: NPI-10 item – forest plot

	Cholinesterase inhibitor							Mean Difference	Mean Difference		
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Fixed, 95% Cl	IV, Fixed, 95% Cl		
1.27.1 Donepezil											
Aarsland 2002	31.8	15.4	12	35.1	8.1	12	41.9%	-3.30 [-13.14, 6.54]			
Ravina 2005 Subtotal (95% CI)	40.3	13.6	21 33	40.5	13.7	20 32	58.1% 100.0 %	-0.20 [-8.56, 8.16] - 1.50 [-7.87, 4.87]			
Heterogeneity: Chi ² = Test for overall effect: .	0.22, df = 1 (Z = 0.46 (P =	(P = 0.64); = 0.64)	I ² = 0%								
Total (95% CI) Heterogeneity: Chi ² = I Test for overall effect: . Test for subgroup diffe	33 ² = 0% e			32	100.0%	-1.50 [-7.87, 4.87] -	-20 -10 0 10 20 Favours medication Favours placebo				

PDD – cholinesterase inhibitor vs placebo: UPDRS III – forest plot

Parkinsons disease dementia – memantine

PDD – memantine vs. placebo: adverse events

		Qualit	y assessment			No of patients			Quality		
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Memantine	Placebo	Relative (95% CI)	Absolute (95% Cl)	Quality	
Any adverse e	vents (p	robability of	experiencing ≥	1; follow-up 1	6 to 24 weeks,	lower is bet	ter); see F	igure 14 for forest plot			
2 ^{1,2}	RCT	not serious	not serious	not serious	serious ³	34/73 (46.6%)	35/72 (48.6%)	RR 0.97 (0.69 to 1.37)	15 fewer per 1000 (from 151 fewer to 180 more)	⊕⊕⊕O MODERATE	
Serious adverse events (probability of experiencing ≥1; follow-up 16 to 24 weeks, lower is better); see Figure 15 for forest plot											
2 ^{1,2}	RCT	not serious	not serious	not serious	very serious3,4	9/73 (12.3%)	8/72 (11.1%)	RR 1.09 (0.45 to 2.67)	10 more per 1000 (from 61 fewer to 186 more)	⊕⊕OO LOW	
Adverse event	s requir	ing treatment	t withdrawal (pr	obability of ex	periencing; fo	llow-up 24 w	veeks, low	ver is better)			
1 ¹	RCT	not serious	N/A	not serious	very serious3,4	6/62 (9.7%)	5/58 (8.6%)	RR 1.12 (0.36 to 3.48)	10 more per 1000 (from 55 fewer to 214 more)	⊕⊕OO LOW	
Emre 2010; data reported for PDD population only; study also included people with DLB Leroi 2009; not clear if adverse event data reported at end of active treatment (16 weeks) or end of drug withdrawal phase (22 weeks) At a 95% confidence level, data are consistent with appreciable harm, appreciable benefit or no difference Very small numbers of events											

	Meman	tine	Place	bo		Risk Ratio	Risk Ratio	
Study or Subgroup	Events	Total	Events	Total	Weight	M-H, Fixed, 95% Cl	M-H, Fixed, 95% Cl	
Emre 2010	28	62	26	58	77.2%	1.01 [0.68, 1.50]		
Leroi 2009	. 6	11	9	14	22.8%	0.85 [0.44, 1.65]		
Total (95% CI)		73		72	100.0%	0.97 [0.69, 1.37]	+	
Total events	34		35					
Heterogeneity: Chi ² = Test for overall effect:	0.19, df = 7 = 0.17	1 (P = P = 0.8	0.66); I ^z = 7)	:0%				
rootior oronali olioot.	- 0.11,	. 0.0	.,				Favours medication Favours placebo	

PDD – memantine vs placebo: any adverse events (proportion of participants experiencing ≥1) – forest plot

	Meman	tine	Place	bo		Risk Ratio	Risk Ratio	
Study or Subgroup	Events	Total	Events	Total	Weight	M-H, Fixed, 95% Cl	M-H, Fixed, 95% Cl	
Emre 2010	8	62	7	58	89.2%	1.07 [0.41, 2.76]		
Leroi 2009	1	11	1	14	10.8%	1.27 [0.09, 18.14]		
Total (95% CI)		73		72	100.0 %	1.09 [0.45, 2.67]		
Total events	9		8					
Heterogeneity: Chi ² =	0.01, df=	1 (P = 1	0.90); I ^z =					
Test for overall effect:	Z = 0.19 (P = 0.8	5)				Favours medication Favours placebo	20

PDD – memantine vs placebo: serious adverse events (proportion of participants experiencing ≥1) – forest plot

PDD – memantine vs. placebo: cognitive function

		Qual	ity assessment			No of pat	tients	Effect	Quality		
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Memantine	Placebo	Mean difference (95% Cl)	Quanty		
MMSE (follow-up	MSE (follow-up 16 weeks; range of scores: 0-30; higher is better)										
1 ¹	RCT	not serious	N/A	not serious	very serious ^{2,3}	10	14	1 lower (6.01 lower to 4.01 higher)	⊕⊕OO LOW		
Clock drawing te	st (follow-	up 24 weeks; ra	nge of scores: 0-1	0; higher is bette	er)						
14	RCT	not serious	N/A	not serious	serious ²	57	56	3.1 higher (6.94 lower to 13.14 higher)	⊕⊕⊕O MODERATE		
¹ Leroi 2009; da: ² At a 95% confi ³ Very small nun ⁴ Emre 2010; da	MODERATE Leroi 2009; data reported for end of drug treatment phase (16 weeks) At a 95% confidence level, data are consistent with appreciable benefit, appreciable harm or no difference Very small numbers of participants in the study Emre 2010: data reported for PDD population only: study also included people with DLB										

PDD – memantine vs. placebo: global assessment

		Quali	ity assessment			No of pa	tients		Quality		
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Memantine	Placebo	Effect (95% CI)	Quality		
ADCS-CGIC (follow-up 24 weeks; range of scores: 1-7; lower is better)											
1 ¹	RCT	not serious	N/A	not serious	serious ²	60	56	MD 0.2 lower (0.69 lower to 0.29 higher)	⊕⊕⊕O MODERATE		
CIBIC+ (at least	minimal	improvement; f	follow-up 16 weel	ks; higher is bet	ter)						
1 ³	RCT	not serious	N/A	not serious	very serious2,4	6/10 (60%)	6/14 (42.9%)	RR 1.4 (0.64 to 3.08) 171 more per 1000 (from 154 fewer to 891 more)	⊕⊕OO LOW		
¹ Emre 2010; c ² At a 95% con ³ Leroi 2009; d ⁴ Data from a s	^(60%) (42.9%) 171 more per 1000 (from 154 fewer to 891 more) LOW ¹ Emre 2010; data reported for PDD population only; study also included people with DLB ² At a 95% confidence level, data are consistent with appreciable benefit, appreciable harm or no difference ³ Leroi 2009; data reported for end of drug treatment phase (16 weeks) ⁴ Data from a single very small study										

PDD - memantine vs. placebo: activities of daily living

		Qualit	ty assessment			No of pat	tients	Effect	Quality	
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Memantine	Placebo	Mean difference (95% CI)	Quanty	
ADCS-ADL (follow-up 24 weeks; measured with: 23-item score; higher is better)										
1 ¹	RCT	not serious	N/A	not serious	serious ²	60	56	0.8 higher (3.22 lower to 4.82 higher)	⊕⊕⊕O MODERATE	
¹ Emre 2010; dat	ta reporte	d for PDD popu	lation only; study a	also included pe	ople with DLB					

² At a 95% confidence level, data are consistent with appreciable benefit, appreciable harm or no difference

PDD – memantine vs. placebo: carer-reported outcomes

		Qu	ality assessment			No of patients Effect		Effect	Quality		
No of studies	Design	Risk of bia	s Inconsistency	Indirectness	Imprecision	Memantine	Placebo	Mean difference (95% CI)	Quanty		
ZBI (follow-up 16	to 24 weel	ks; lower is b	etter) ¹ ; see Figure 16	for forest plot							
2 ^{2,3}	RCT	not serious	not serious	not serious	serious⁴	71	70	3.4 lower (7.21 lower to 0.42 higher)	⊕⊕⊕O MODERATE		
 ¹ Data from Leron ² Leroi 2009; data ³ Emre 2010; data ⁴ At a 95% confide 	Data from Leroi 2009 reported in a secondary publication (Leroi 2014) Leroi 2009; data reported for end of drug treatment phase (16 weeks) Emre 2010; data reported for PDD population only; study also included people with DLB At a 95% confidence level, data are consistent with appreciable benefit, appreciable harm or no difference										
Study or Subgro	up Mea	Memantine an SD	Placebo Total Mean SE) Total Weight	Mean Difference IV, Fixed, 95%	e Cl	Mean IV, Fix	ı Difference xed, 95% Cl			



PDD – memantine vs placebo: ZBI – forest plot

PDD – m	emantine vs.	placebo:	other	non-cognitive	outcomes
		p1000000	•••••		

		Qual	ity assessment			No of pat	tients	Effect	Quality		
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Memantine	Placebo	Mean difference (95% CI)			
IPI 12-item (follow-up 24 weeks; range of scores: 0-144; lower is better)											
1 ¹	RCT	not serious	N/A	not serious	serious ³	60	56	MD 1.50 lower (6.35 lower to 3.35 higher)	⊕⊕⊕O MODERATE		
NPI 10-item (follo	IPI 10-item (follow-up 16 weeks; range of scores: 0-120; lower is better)										
1 ²	RCT	not serious	N/A	not serious	very serious3,4	10	14	MD 2.00 lower (11.64 lower to 7.64 higher)	⊕⊕OO LOW		
UPDRS III (follow	/-up 16 to	24 weeks; lowe	er is better); see Fi	gure 17 for fores	st plot						
2 ^{1,2}	RCT	not serious	not serious	not serious	serious ^{3,5}	70	70	MD 0.88 higher (2.35 lower to 4.1 higher)	⊕⊕⊕O MODERATE		
 ¹ Emre 2010; da ² Leroi 2009; da ³ At a 95% cont ⁴ Data from a si 	Emre 2010; data reported for PDD population only; study also included people with DLB MODERATE Leroi 2009; data reported for end of drug treatment phase (16 weeks) At a 95% confidence level, data are consistent with appreciable benefit, appreciable harm or no difference Data from a single very small study Study										

^oCI cross MID between 3.25 (Horvath et al 2015) and 5 points (Schrag et al., 2006)

	Memantine				Placebo			Mean Difference	Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Fixed, 95% Cl	IV, Fixed, 95% CI
Emre 2010	1.5	9.6776	60	1	10.0821	56	80.2%	0.50 [-3.10, 4.10]	
Leroi 2009	24.3	8.8	10	21.9	9.1	14	19.8%	2.40 [-4.84, 9.64]	-
Total (95% CI)			70			70	100.0%	0.88 [-2.35, 4.10]	• • • •
Heterogeneity: Chi ² = 0.21, df = 1 (P = 0.65); l ² = 0%									

PDD – memantine vs placebo: UPDRS III – forest plot

Dementia with Lewy bodies – cholinesterase inhibitors

DLB - cholinesterase inhibitor vs. placebo: adverse events

Quality assessment	No of patients		Quality					
No of studies Design Risk of bias Inconsistency Indirectness Imprecision	Chl Placebo	Relative (95% CI)	Relative (95% CI) Absolute (95% CI)					
Any adverse events – cholinesterase inhibitors (probability of experiencing ≥1; follow-up 12 to 20 weeks);); see Figure 18 for forest plot								

		Qualit	y assessment			No of p	patients		Quality	
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Chl	Placebo	Relative (95% CI)	Absolute (95% Cl)	Quanty
3 ^{1–3}	RCT	not serious	not serious	not serious	serious⁴	201/260 (77.3%)	101/141 (71.6%)	RR 1.11 (0.98 to 1.25)	79 more per 1000 (from 14 fewer to 179 more)	⊕⊕⊕O MODERATE
Any adverse ev	vents – d	donepezil (pro	obability of expe	riencing ≥1; fo	llow-up 12 we	eks)				
2 ^{1,2}	RCT	not serious	not serious	not serious	serious⁴	147/201 (73.1%)	55/80 (68.8%)	RR 1.05 (0.88 to 1.25)	34 more per 1000 (from 83 fewer to 172 more)	⊕⊕⊕O MODERATE
Any adverse ev	vents – I	rivastigmine (probability of ex	periencing ≥1	; follow-up 20	weeks)				
1 ³	RCT	not serious	N/A	not serious	not serious	54/59 (91.5%)	46/61 (75.4%)	RR 1.21 (1.03 to 1.43)	158 more per 1000 (from 23 more to 324 more)	⊕⊕⊕⊕ HIGH
Serious advers	se events	s – cholineste	erase inhibitors (probability of	experiencing	≥1; follov	v-up 12 to	20 weeks);); see Figur	e 19 for forest plot	
3 ^{1–3}	RCT	not serious	not serious	not serious	serious ⁴	23/260 (8.8%)	15/141 (10.9%)	RR 0.98 (0.53 to 1.82)	2 fewer per 1000 (from 51 fewer to 89 more)	⊕⊕⊕O MODERATE
Serious advers	se events	s – donepezil	(probability of e	xperiencing ≥	1; follow-up 12	2 weeks)				
2 ^{1,2}	RCT	not serious	not serious	not serious	serious⁴	13/201 (6.5%)	7/80 (8.8%)	RR 0.73 (0.3 to 1.81)	24 fewer per 1000 (from 61 fewer to 71 more)	⊕⊕⊕O MODERATE
Serious advers	se events	s – rivastigmi	ne (probability o	of experiencing	g ≥1; follow-up	o 20 week	(S)			
1 ³	RCT	not serious	N/A	not serious	serious⁴	10/59 (16.9%)	8/61 (13.1%)	RR 1.29 (0.55 to 3.05)	38 more per 1000 (from 59 fewer to 269 more)	⊕⊕⊕O MODERATE
Adverse events	s requiri	ing treatment	withdrawal - ch	olinesterase ir	hibitors (prol	bability o	f experien	cing; follow-up 12 to 20	weeks)); see Figure 20 for forest plot	
3 ^{1–3}	RCT	not serious	not serious	not serious	serious⁴	25/260 (9.6%)	16/141 (11.3%)	RR 0.9 (0.49 to 1.63)	11 fewer per 1000 (from 58 fewer to 71 more)	⊕⊕⊕O MODERATE
Adverse events	s requiri	ing treatment	withdrawal - do	nepezil (proba	ability of exper	riencing;	follow-up	12 weeks)		
2 ^{1,2}	RCT	not serious	not serious	not serious	serious⁴	18/201 (9%)	9/80 (11.3%)	RR 0.82 (0.39 to 1.74)	20 fewer per 1000 (from 69 fewer to 83 more)	⊕⊕⊕O MODERATE
Adverse events	s requiri	ing treatment	withdrawal - riv	astigmine (pro	bability of ex	periencin	g; follow-	up 20 weeks)		
1 ³	RCT	not serious	N/A	not serious	serious ⁴	7/59 (11.9%)	7/61 (11.5%)	RR 1.03 (0.39 to 2.77)	3 more per 1000 (from 70 fewer to 203 more)	⊕⊕⊕O MODERATE
¹ Ikeda 2015; ² Mori 2012; d ³ McKeith 200	data for lata for : 0	2 active trea 3 active treat	atment groups w ment groups we	vere combine ere combined	d (donepezil (donepezil 3i	5mg and mg, 5mg	10mg) and 10m	ng)		

⁴ At a 95% confidence level, data are consistent with appreciable harm, appreciable benefit or no difference
	Cholinesterase inhibitor			bo		Risk Ratio	Risk Ratio		
Study or Subgroup	Events	Total	Events	Total	Weight	M-H, Fixed, 95% Cl	M-H, Fixed, 95% Cl		
4.1.1 Donepezil									
lkeda 2015	64	96	31	46	34.0%	0.99 [0.77, 1.26]	+		
Mori 2012	83	105	24	34	29.4%	1.12 [0.88, 1.42]			
Subtotal (95% CI)		201		80	63.3%	1.05 [0.88, 1.25]	◆		
Total events	147		55						
Heterogeneity: Chi ² =	0.51, df = 1 (P = 0.48); I² = 0 9	6						
Test for overall effect	Z = 0.56 (P = 0.58)								
4.1.2 Rivastigmine									
McKeith 2000	54	59	46	61	36.7%	1.21 [1.03, 1.43]	-		
Subtotal (95% CI)		59		61	36.7%	1.21 [1.03, 1.43]	◆		
Total events	54		46						
Heterogeneity: Not ap	oplicable								
Test for overall effect	: Z = 2.33 (P = 0.02)								
Total (95% CI)		260		141	100.0%	1.11 [0.98, 1.25]	•		
Total events	201		101						
Heterogeneity: Chi ^z =	2.00, df = 2 (P = 0.37); I ² = 09	6						
Test for overall effect:	Z = 1.67 (P = 0.10)						U.UD U.Z 1 5 ZU		
Test for subgroup dif	ferences: Chi ² = 1.44,	df = 1 (l	P = 0.23).	. I ² = 30	.8%		ravours medication ravours placebo		

DLB – cholinesterase inhibitor vs placebo: any adverse events (proportion of participants experiencing ≥1) – forest plot

	Cholinesterase inhibitor			bo		Risk Ratio	Risk Ratio
Study or Subgroup	Events	Total	Events	Total	Weight	M-H, Fixed, 95% Cl	M-H, Fixed, 95% Cl
4.2.1 Donepezil							
lkeda 2015	5	96	5	46	38.3%	0.48 [0.15, 1.57]	_
Mori 2012	8	105	2	34	17.1%	1.30 [0.29, 5.81]	
Subtotal (95% CI)		201		80	55.4%	0.73 [0.30, 1.81]	
Total events	13		7				
Heterogeneity: Chi ² =	1.04, df = 1 (P = 0.31)); I ² = 49	6				
Test for overall effect:	Z = 0.68 (P = 0.50)						
4.2.2 Rivastigmine							
McKeith 2000	10	59	8	61	44.6%	1.29 [0.55, 3.05]	
Subtotal (95% CI)		59		61	44.6%	1.29 [0.55, 3.05]	
Total events	10		8				
Heterogeneity: Not ap	oplicable						
Test for overall effect:	Z = 0.59 (P = 0.56)						
Total (95% CI)		260		141	100.0%	0.98 [0.53, 1.82]	-
Total events	23		15				
Heterogeneity: Chi ² =	1.92, df = 2 (P = 0.38)); I = 09	6				
Test for overall effect:	Z = 0.06 (P = 0.95)	-					U.U5 U.2 1 5 2U
Test for subgroup dif	ferences: Chi ² = 0.80,	df = 1 (i	P = 0.37),	.l ² = 09	6		Favours medication Favours placebo

DLB – cholinesterase inhibitor vs placebo: serious adverse events (proportion of participants experiencing ≥1) – forest plot



DLB – cholinesterase inhibitor vs placebo: adverse events requiring treatment withdrawal (proportion of participants experiencing) – forest plot

DLB - cholinesterase inhibitor vs. placebo: cognitive function

		Qualit	ty assessment			No	of patients	Effect	Quality			
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Chl	Placebo	Mean difference (95% CI)	Quality			
MMSE – cholinest	erase inhil	pitors (follow-up	12 to 20 weeks; ran	ge of scores: 0-30	0; higher is better); see	Figure 21 for	forest plot				
3 ^{1–3}	RCT	not serious	serious ⁴	not serious	not serious	256	136	1.77 higher (1.06 to 2.47 higher)	⊕⊕⊕O MODERATE			
MMSE – donepezi	l (follow-u	o 12 weeks; rang	e of scores: 0-30; h	igher is better)								
2 ^{1,3}	RCT	not serious	serious ⁴	not serious	not serious	197	75	1.91 higher (1.11 to 2.71 higher)	⊕⊕⊕O MODERATE			
MMSE – rivastigm	ine (follow	-up 20 weeks; ra	inge of scores: 0-30	; higher is better)								
1²	RCT	not serious	N/A	not serious	serious⁵	59	61	1.24 higher (0.28 lower to 2.76 higher)	⊕⊕⊕O MODERATE			
¹ Ikeda 2015; dat ² McKeith 2000; d	keda 2015; data for 2 active treatment groups were combined (donepezil 5mg and 10mg) AcKeith 2000; data for this outcome taken from a Cochrane review; data not reported in published paper											

³ Mori 2012; data for 3 active treatment groups were combined (donepezil 3mg, 5mg and 10mg)

⁴ i² >40% between studies

⁵ At a 95% confidence level, data are consistent with appreciable benefit, appreciable harm or no difference



DLB - cholinesterase inhibitor vs placebo: MMSE - forest plot

DLB – cholinesterase inhibitor vs. placebo: global assessment

				,								
		Qualit	ty assessment			No of	patients		Quality			
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Chl	Placebo	Ellect (95% Cl)	Quality			
CIBIC+ – donepez	zil (follow-	up 12 weeks; ra	nge of scores: 1-7;	lower is better) ¹	l i							
1 ²	RCT	not serious	N/A	not serious	not serious	91	30	MD 1.17 lower (1.66 to 0.68 lower)	⊕⊕⊕⊕ HIGH			
CIBIC+ – donepez	zil (at leas	t minimal improv	vement; follow-up	12 weeks; higher	is better)							
1 ²	RCT	not serious	N/A	not serious	not serious	62/91 (68.1%)	10/30 (33.3%)	RR 2.04 (1.21 to 3.46) 347 more per 1000 (from 70 more to 820 more)	⊕⊕⊕⊕ HIGH			
¹ Mean and SD of ² Mori 2012: data	lean and SD calculated from data presented in paper lori 2012: data for 3 active treatment droups were combined (donepezil 3mg, 5mg and 10mg)											

DLB – cholinesterase inhibitor vs. placebo: carer-reported outcomes

		Quali	ity assessment		No	of patients	Effect		
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Chl	Placebo	Mean difference (95% Cl)	Quality
ZBI - donepezil (folle	ow-up 12 w	eeks; lower is bett	er); see Figure 22 for	forest plot					
2 ^{1,2}	RCT	not serious	not serious	not serious	not serious	191	77	4.49 lower (7.64 to 1.34 lower)	⊕⊕⊕⊕ HIGH
¹ Ikeda 2015; data ² Mori 2012; data f	for 2 active or 3 active	e treatment group treatment groups	s were combined (do were combined (dor	nepezil 5mg and nepezil 3mg, 5mg a	10mg) and 10mg)				

	Choline	sterase inh	se inhibitor Placebo					Mean Difference	Mean Difference		
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Fixed, 95% Cl	IV, Fixed, 95% Cl		
4.7.1 Donepezil											
lkeda 2015	-2.8562	12.2368	96	-0.1	12.21	46	53.8%	-2.76 [-7.05, 1.54]			
Mori 2012	-2.3116	14.1638	95	4.2	10.4	31	46.2%	-6.51 [-11.15, -1.87]			
Subtotal (95% CI)			191			77	100.0%	-4.49 [-7.64, -1.34]	◆		
Heterogeneity: Chi ² =	1.36, df = 1	1 (P = 0.24);	l ² = 26%								
Test for overall effect:	Z = 2.79 (F	P = 0.005)									
Total (95% CI)			101			77	100.0%	4 49 [7 64 1 34]			
10(a) (35% C)			191				100.0%	-4.49 [-7.04, -1.54]			
Heterogeneity: Chi ² =	1.36, df = 1	1 (P = 0.24);	; l² = 26%					-	-20 -10 0 10 20		
Test for overall effect: .	Z = 2.79 (F	P = 0.005)							Eavours medication Eavours placebo		
Test for subgroup diffe	erences: N	lot applicab	le						ravours medication ravours placebo		

DLB – cholinesterase inhibitor (donepezil) vs placebo: ZBI – forest plot

DLB – cholinesterase inhibitor vs. placebo: Other non-cognitive outcomes

		Quality	assessment		No of p	atients	Effect	Quality	
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Chl	placebo	Mean difference (95% CI)	
NPI-10 item – choline	esterase inh	ibitors (follow-up 1	2 to 20 weeks; range of	of scores: 0-120; low	ver is better) ¹	see Figu	re 23 for	forest plot	
32-4	RCT	not serious	serious⁵	not serious	serious ⁶	243	129	2.06 lower (7.15 lower to 3.02 higher)	⊕⊕OO LOW
NPI-10 item – donepo	ezil (follow-	up 12 weeks; range	of scores: 0-120; lowe	er is better) ¹					
2 ^{2,4}	RCT	not serious	serious⁵	not serious	serious ⁶	196	76	1.54 lower (9.37 lower to 6.29 higher)	⊕⊕OO LOW
NPI-10 item – rivastig	gmine (follo	w-up 20 weeks; rar	nge of scores: 0-120; lo	ower is better)					
1 ³	RCT	not serious	N/A	not serious	serious ⁶	47	53	3.8 lower (9.25 lower to 1.65 higher)	⊕⊕⊕O MODERATE
NPI-4 item – cholines	sterase inhil	bitors (follow-up 12	to 20 weeks; range of	scores: 0-48; lower	is better)7; s	ee Figure	e 24 for fo	rest plot	
23,4	RCT	not serious	not serious	not serious	not serious	161	93	2.49 lower (4.64 to 0.33 lower)	⊕⊕⊕⊕ HIGH
NPI-4 item – donepez	zil (follow-u	o 12 weeks; range o	of scores: 0-48; lower i	s better) ⁷					
14	RCT	not serious	N/A	not serious	not serious	102	32	3.59 lower (6.93 to 0.25 lower)	⊕⊕⊕⊕ HIGH
NPI-4 item – rivastigi	mine (follow	up 20 weeks; rang	je of scores: 0-48; low	er is better) ⁷					
1 ³	RCT	not serious	N/A	not serious	serious ⁶	59	61	1.7 lower (4.52 lower to 1.12 higher)	⊕⊕⊕O MODERATE
NPI-2 item – donepez	il (follow-u	o 12 weeks; range o	of scores: 0-24; lower i	s better)8; see Figur	e 25 for fores	t plot			
22,4	RCT	not serious	serious⁵	not serious	serious ⁶	196	76	2.3 lower (6.32 lower to 1.72 higher)	⊕⊕OO LOW
UPDRS III – cholines	terase inhib	itors (follow-up 12	weeks; lower is better) ¹ ; see Figure 26 for	forest plot				
22,4	RCT	serious ⁹	not serious	not serious	not serious ¹⁰	195	77	0.67 lower (2.08 lower to 0.73 higher)	⊕⊕⊕O MODERATE
UPDRS III – donepez	il (follow-up	12 weeks; lower is	s better) ¹						
2 ^{2,4}	RCT	not serious	not serious	not serious	not serious ¹⁰	195	77	0.67 lower (2.08 lower to 0.73 higher)	⊕⊕⊕⊕ HIGH
 ¹ SD not reported fc ² Ikeda 2015; data f ³ McKeith 2000 ⁴ Mori 2012; data fo ⁵ i² >40% between s ⁶ At a 95% confiden ⁷ NPI 4-item consist ⁸ NPI 2-item consist ⁹ Data for outcome ¹⁰ CI do not cross M 	or this outco for 2 active r 3 active to studies ce level, da ts of 4 NPI ts of 2 NPI not present 11D betweet	ome in Ikeda 2015 treatment groups w ata are consistent domains – hallucin domains – hallucin domains – hallucin ted in McKeith 200 n 3.25 (Horvath et	; calculated from SE i were combined (done were combined (done with appreciable bene nations, delusions, dy nations and cognitive 00. Study reported no al., 2015) and 5 poin	reported in paper apezil 5mg and 10m pezil 3mg, 5mg and afit, appreciable hau sphoria and apathy fluctuation significant difference ts (Schrag et al., 20	ng) 10mg) m or no diffe ce between <u>c</u> 006)	rence troups			

	Choline	sterase inhi	bitor	PI	acebo			Mean Difference		Mean Difference	
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Random, 95% Cl		IV, Random, 95% Cl	
4.8.1 Donepezil											
lkeda 2015	-4.4468	9.618	94	-6.4	9.95	44	41.0%	1.95 [-1.57, 5.48]		- +	
Mori 2012	-5.8088	15.3238	102	0.3	17.5	32	26.9%	-6.11 [-12.86, 0.64]			
Subtotal (95% CI)			196			76	67.8%	-1.54 [-9.37, 6.29]			
Heterogeneity: Tau ² =	24.94; Ch	i ^z = 4.30, df =	= 1 (P = I	0.04); I ^z	= 77%						
Test for overall effect:	Z = 0.39 (F	P = 0.70)									
4.8.2 Rivastigmine											
McKeith 2000	-5	16.2	47	-1.2	10.7	53	32.2%	-3.80 [-9.25, 1.65]			
Subtotal (95% CI)			47			53	32.2%	-3.80 [-9.25, 1.65]			
Heterogeneity: Not ap	plicable										
Test for overall effect:	Z=1.37 (F	P = 0.17)									
Total (95% CI)			243			129	100.0%	-2.06 [-7.15, 3.02]			
Heterogeneity: Tau ² =	: 13.19: Ch	i ^z = 5.88. df :	= 2 (P =	0.05): P	= 66%				—		
Test for overall effect:	Z = 0.79 (F	P = 0.43	- 0						-20	-10 0 10	20
Test for subaroup diff	erences: C	;hi ² = 0.22. d	lf = 1 (P :	= 0.64).	l ² = 0%	6				Favours medication Favours placebo	

DLB – cholinesterase inhibitor vs placebo: NPI-10 item – forest plot

	Cholines	terase inhi	ibitor	Pla	Placebo Mean Differer			Mean Difference	Mean Difference				
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Fixed, 95% CI		IV, Fix	ed, 95% Cl		
4.9.1 Donepezil													
Mori 2012 Subtotal (95% CI)	-3.8912	8.1471	102 102	-0.3	8.5	32 32	41.6% 41.6 %	-3.59 [-6.93, -0.25] - 3.59 [-6.93, -0.25]		-	•		
Heterogeneity: Not ap	plicable												
Test for overall effect:	Z= 2.11 (P	= 0.04)											
4.9.2 Rivastigmine													
McKeith 2000	-2.5	8.4	59	-0.8	7.3	61	58.4%	-1.70 [-4.52, 1.12]			┡╋		
Subtotal (95% CI)			59			61	58.4 %	-1.70 [-4.52, 1.12]					
Heterogeneity: Not ap	plicable												
Test for overall effect:	Z=1.18 (P	= 0.24)											
Total (95% CI)			161			93	100.0%	-2.49 [-4.64, -0.33]		-	•		
Heterogeneity: Chi ² =	0.72, df = 1	(P = 0.40);	I²=0%						H		<u> </u>	-+	
Test for overall effect:	Z = 2.26 (P	= 0.02)							-20	-10 Eavoure modicatio	u n Eavour	10 s placobo	20
Test for subgroup diff	erences: Cl	hi² = 0.72, c	if = 1 (P =	= 0.40),	I ² = 0	%				ravours medicatio	n ravour	s placebu	

DLB – cholinesterase inhibitor vs placebo: NPI-4 item – forest plot

	Cholines	terase inhi	ibitor	Pla	acebo)		Mean Difference	Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Random, 95% Cl	IV, Random, 95% Cl
4.10.1 Donepezil									
lkeda 2015	-2.3255	4.5291	94	-2	4.2	44	51.9%	-0.33 [-1.87, 1.22]	
Mori 2012 Subtotal (95% Cl)	-3.3343	5.0689	102 196	1.1	5.7	32 76	48.1% 100.0 %	-4.43 [-6.64, -2.23] - 2.30 [-6.32, 1.72]	
Heterogeneity: Tau ² : Test for overall effect	= 7.50; Chiᢪ∘ t: Z = 1.12 (P	= 8.95, df = = 0.26)	1 (P = 0.	.003); I²	= 899	%			
Total (95% CI)			196			76	100.0%	-2.30 [-6.32, 1.72]	-
Heterogeneity: Tau ^z : Test for overall effect Test for subgroup dit	= 7.50; Chi <mark>²</mark> ÷ t: Z = 1.12 (P fferences: N	= 8.95, df = = 0.26) ot applicab	1 (P = 0. le	.003); I²	= 899	%			-20 -10 0 10 20 Favours medication Favours placebo

DLB - cholinesterase inhibitor (donepezil) vs placebo: NPI-2 item - forest plot



DLB – cholinesterase inhibitor (donepezil) vs placebo: UPDRS III – forest plot

Dementia with Lewy bodies – memantine

DLB - memantine vs. placebo: adverse events

	Qualit	y assessment			No of pa	tients		Effect	Quality		
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Memantine	Placebo	Relative (95% CI)	Absolute (95% CI)	Quanty	
Any adverse e	vents (p	probability of	experiencing ≥ [,]	1; follow-up 24	ł weeks)						
1 ¹	RCT	not serious	N/A	not serious	serious ²	18/34 (52.9%)	17/41 (41.5%)	RR 1.28 (0.79 to 2.07)	116 more per 1000 (from 87 fewer to 444 more)	⊕⊕⊕O MODERATE	
Serious adver	se event	ts (probability	of experiencin	g ≥1; follow-u	p 24 weeks)						
1 ¹	RCT	not serious	N/A	not serious	very serious2,3	6/34 (17.6%)	3/41 (7.3%)	RR 2.41 (0.65 to 8.93)	103 more per 1000 (from 26 fewer to 580 more)	⊕⊕OO LOW	
Adverse event	s requir	ing treatment	t withdrawal (pr	obability of ex	(periencing; fo	llow-up 24 v	/eeks)				
1 ¹	RCT	not serious	N/A	not serious	very serious2,3	5/34 (14.7%)	7/41 (17.1%)	RR 0.86 (0.3 to 2.47)	24 fewer per 1000 (from 120 fewer to 251 more)	⊕⊕OO LOW	
¹ Emre 2010; data reported for DLB population only; study also included people with PDD											

² At a 95% confidence level, data are consistent with appreciable harm, appreciable benefit or no difference

³ Very small numbers of events

DLB - memantine vs. placebo: cognitive outcomes

		-	_									
		Qualit	y assessment			No of pat	tients	Effect	Quality			
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Memantine	Placebo	Mean difference (95% Cl)				
Clock drawing tes	st (follow-	up 24 weeks; ran	ge of scores: 0-10	; higher is better)	ĺ							
1 ¹	RCT	not serious	N/A	not serious	serious ²	33	43	1.3 higher (0.51 lower to 3.11 higher)	⊕⊕⊕O MODERATE			
¹ Emre 2010; da	Emre 2010; data reported for DLB population only; study also included people with PDD											
2 4 4 4 4 4 4												

² At a 95% confidence level, data are consistent with appreciable benefit, appreciable harm or no difference

DLB – memantine vs. placebo: global assessment

		Quali	ty assessment			No of pat	tients	Effect	Quality			
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Memantine	Placebo	Mean difference (95% CI)	Quanty			
ADCS-CGIC (follo	DCS-CGIC (follow-up 24 weeks; lower is better)											
1 ¹	RCT	not serious	N/A	not serious	serious ²	33	41	0.6 lower (1.22 lower to 0.02 higher)	⊕⊕⊕O MODERATE			
¹ Emre 2010; da	ta reporte	d for DLB popu	lation only; study a	also included pe	ople with PDD							

² At a 95% confidence level, data are consistent with appreciable benefit, appreciable harm or no difference

DLB - memantine vs. placebo: activities of daily living

		Quali	ty assessment			No of pat	tients	Effect	Quality			
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Memantine	Placebo	Mean difference (95% CI)	Quanty			
ADCS-ADL (follow	CS-ADL (follow-up 24 weeks; range of scores: 0-78; higher is better)											
1 ¹	RCT	not serious	N/A	not serious	serious ²	33	41	1.6 higher (4.9 lower to 8.1 higher)	⊕⊕⊕O MODERATE			
¹ Emre 2010; dat ² Wide 95% cont	ta reporte idence int	d for DLB popul ervals. data are	ation only; study a consistent with an	lso included peo preciable benefi	ple with PDD it. appreciable h	arm or no diffe	rence					

DLB – memantine vs. placebo: carer-reported outcomes

		Quali	ty assessment			No of pat	tients	Effect	Quality			
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Memantine	Placebo	Mean difference (95% CI)	Quanty			
ZBI (follow-up 24	I (follow-up 24 weeks; lower is better)											
1 ¹	RCT	not serious	N/A	not serious	serious ²	33	41	1.4 lower (6.66 lower to 3.86 higher)	⊕⊕⊕O MODERATE			
¹ Emre 2010; dat	ta reporte	d for DLB popu	lation only; study a	also included peo	ople with PDD							

² Wide 95% confidence intervals, data are consistent with appreciable benefit, appreciable harm or no difference

DLB – memantine vs. placebo: other non-cognitive outcomes

			U								
		Quali	ty assessment			No of pat	tients	Effect	Quality		
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Memantine	Placebo	Mean difference (95% Cl)	Quanty		
VPI-12 item (follow-up 24 weeks; range of scores: 0-144; lower is better)											
1 ¹	RCT	not serious	N/A	not serious	serious ²	33	41	6 lower (12.23 lower to 0.23 higher)	⊕⊕⊕O MODERATE		
UPDRS III (follow-	-up 24 wee	ks; lower is bett	ter)								
1 ¹	RCT	not serious	N/A	not serious	serious ^{2,3}	33	41	1.4 lower (5.52 lower to 2.72 higher)	⊕⊕⊕O MODERATE		
¹ Emre 2010; da	ta reporte	d for DLB popu	lation only; study a	also included peo	ople with PDD						
² Wide 95% con	fidence in	tervals, data are	e consistent with a	ppreciable bene	fit, appreciable l	narm or no diffe	erence				
³ CI cross the MI	D betweel	n 3.25 (Horvath	et al., 2015) and 5	5 points (Schrag	et al., 2006)						

Mixed population (PDD or DLB) – cholinesterase inhibitors

PDD/DLB - cholinesterase inhibitor vs. placebo: adverse events

		Qualit	y assessment			No of p	atients		Quality	
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Chl	Placebo	Relative (95% CI)	Absolute (95% Cl)	Quanty
Any adverse ev	vents – o	cholinesteras	e inhibitors (pro	bability of exp	eriencing ≥1	; follow-up	o 10 to 24	weeks; lower is better);	see Figure 27 for forest plot	
7 ^{1–7}	RCT	not serious	not serious	not serious	not serious	810/1034 (78.3%)	369/525 (70.3%)	RR 1.12 (1.05 to 1.19)	84 more per 1000 (from 35 more to 134 more)	⊕⊕⊕⊕ HIGH
Any adverse ev	vents – o	donepezil (pro	obability of expe	riencing ≥1; f	follow-up 10 to	o 24 weeks	s; lower is	better)		
5 ^{1,2,4,6,7}	RCT	not serious	not serious	not serious	serious ⁸	453/613 (73.9%)	196/285 (68.8%)	RR 1.06 (0.97 to 1.16)	41 more per 1000 (from 21 fewer to 110 more)	⊕⊕⊕O MODERATE
Any adverse ev	vents – I	rivastigmine (probability of ex	periencing ≥	1; follow-up 2	0 to 24 we	eks; lowe	r is better)		
2 ^{3,5}	RCT	not serious	not serious	not serious	not serious	357/421 (84.8%)	173/240 (72.1%)	RR 1.19 (1.09 to 1.3)	137 more per 1000 (from 65 more to 216 more)	⊕⊕⊕⊕ HIGH
Serious advers	se events	s – cholineste	erase inhibitors	(probability of	experiencing	≥1; follov	v-up 12 to	24 weeks; lower is bet	ter); see Figure 28 for forest plot	
5 ²⁻⁶	RCT	not serious	not serious	not serious	serious ⁸	137/999 (13.7%)	63/493 (12.8%)	RR 1.10 (0.83 to 1.45)	13 more per 1000 (from 22 fewer to 58 more)	⊕⊕⊕O MODERATE
Serious advers	se events	s – donepezil	(probability of e	xperiencing	≥1; follow-up [•]	12 to 24 w	eeks; low	er is better)		
32,4,6	RCT	not serious	not serious	not serious	serious ⁸	80/578 (13.8%)	29/253 (11.5%)	RR 1.23 (0.83 to 1.84)	26 more per 1000 (from 19 fewer to 96 more)	⊕⊕⊕O MODERATE
Serious advers	se events	s – rivastigmi	ine (probability o	of experiencing	g ≥1; follow-u	ip 20 to 24	weeks; lo	ower is better)		
2 ^{3,5}	RCT	not serious	not serious	not serious	serious ⁸	57/421 (13.5%)	34/240 (14.2%)	RR 0.97 (0.65 to 1.43)	4 fewer per 1000 (from 50 fewer to 61 more)	⊕⊕⊕O MODERATE
Adverse events	s requiri	ing treatment	withdrawal - ch	olinesterase i	nhibitors (pro	bability of	experien	cing; follow-up 10 to 24	weeks; lower is better); see Figure 29 for forest	plot
6 ¹⁻⁶	RCT	not serious	not serious	not serious	not serious	147/1013 (14.5%)	49/505 (9.7%)	RR 1.50 (1.10 to 2.04)	49 more per 1000 (from 10 more to 101 more)	⊕⊕⊕⊕ HIGH

Adverse even	nts requiri	ing treatment	withdrawal –	donepezil	l (proba	bility of e	experiencing; f	ollow-up	10 to 24 weeks; lower is	s better)		
4 1,2,4,6	RCT	not serious	not serious	not seri	ious	serious ⁸	78/592 (13.2%)	28/265 (10.6%)	RR 1.25 (0.84 to 1.87)	26 more per 1000 (from	17 fewer to 92 more)	⊕⊕⊕O MODERATE
Adverse even	nts requiri	ing treatment	withdrawal –	rivastigm	ine (pro	bability	of experiencing	; follow-	up 20 to 24 weeks; lowe	r is better)		
2 ^{3,5}	RCT	not serious	not serious	not seri	ious	not seriou	us 69/421 (16.4%)	21/240 (8.8%)	RR 1.88 (1.17 to 3.03)	77 more per 1000 (from	15 more to 178 more)	⊕⊕⊕⊕ HIGH
¹ Aarsland 20	002											
² Dubois 201	2; data fo	or 2 active tr	eatment group	os were d	combine	ed (done	pezil 5mg and	10mg).	Mean and standard de	viation calculated from o	data reported in papel	r
³ Emre 2004												
⁴ Ikeda 2015; 5 MaKaith 201	; data for	r 2 active trea	atment groups	were co	mbined	d (donep	ezil 5mg and 1	10mg)				
6 Mari 2012:	data for	2 active trea	tmont around		nhinad	(donone	zil 2ma 5ma	and 10m	a)			
⁷ Povino 200	<i>uala 101</i> .	S active tiea	uneni gioups		nomeu	luonepe	zii Sing, Sing e		<i>g)</i>			
* Ravina 200	·5											
° 41 2 45% M	ontidenci	a laval data	are consisten	t with an	nreciah	le henet	fit annreciahle	hamor	no difference			
° At a 95% co	onfidence	e level, data	are consisten	t with app	preciab	le benef	fit, appreciable	harm or	no difference			
° At a 95% co	onfidence	e level, data Cholinester	are consisten ase inhibitor	t with ap	breciab bo	ole benef	fit, appreciable Risk Ratio	harm or	no difference Risk Ra	atio		
• At a 95% cc	onfidence bgroup	e level, data Cholinester Events	are consisten ase inhibitor ; Total	<i>t with ap</i> Placel Events	breciab bo Total	Weight	fit, appreciable Risk Ratio M-H, Fixed, 95%	harm or % Cl	no difference Risk Ra M-H, Fixed,	atio 95% Cl		
• At a 95% cc <u>Study or Sub</u> 6.1.1 PDD	onfidence bgroup	e level, data Cholinester Events	are consisten ase inhibitor ; Total	<i>t with ap</i> Placel Events	breciab bo Total	veight	fit, appreciable Risk Ratio M-H, Fixed, 95%	harm or % Cl	no difference Risk Ra M-H, Fixed,	atio 95% Cl		
• At a 95% cc Study or Sult 6.1.1 PDD Aarsland 200	onfidence bgroup 02	e level, data Cholinester Events 12	are consisten ase inhibitor <u>Total</u>	t with app Placel Events 9	preciab bo <u>Total</u> 12	Weight	Risk Ratio Nisk Ratio M-H, Fixed, 959 1.14 [0.77, 1	ham or <u>6 CI</u> .69]	no difference Risk Ra M-H, Fixed,	atio 95% Cl		
• At a 95% cc <u>Study or Sub</u> 6.1.1 PDD Aarsland 200 Dubois 2012	onfidence bgroup 02 2	e level, data Cholinester Events 12 283	are consisten ase inhibitor <u>Total</u> 14 3 377	<i>t with ap</i> Placel Events 9 123	bo Total 12 173	Weight 2.0% 35.1%	fit, appreciable Risk Ratio <u>M-H, Fixed, 959</u> 1.14 [0.77, 1 1.06 [0.94, 1	ham or <u>% CI</u> .69] .18]	no difference Risk Ra M-H, Fixed,	atio 95% Cl		
• At a 95% cc <u>Study or Sub</u> 6.1.1 PDD Aarsland 200 Dubois 2012 Emre 2004	onfidence bgroup 02 2	e level, data Cholinester Events 12 283 303	are consisten ase inhibitor Total 14 377 3 362	Placel Placel Events 9 123 127	bo Total 12 173 179	Weight 2.0% 35.1% 35.3%	fit, appreciable Risk Ratio <u>M-H, Fixed, 959</u> 1.14 [0.77, 1 1.06 [0.94, 1 1.18 [1.06, 1	harm or <u>6 CI</u> .69] .18] .31]	no difference Risk Ra M-H, Fixed,	atio 95% Cl 		
• At a 95% cc <u>Study or Sut</u> 6.1.1 PDD Aarsland 200 Dubois 2012 Emre 2004 Ravina 2005	onfidence bgroup 02 2	e level, data Cholinester Events 12 283 303 11	are consisten ase inhibitor 7 Total 2 14 3 377 3 362 21	Placel Placel Events 9 123 127 9	bo Total 12 173 179 20	Weight 2.0% 35.1% 35.3% 1.9%	fit, appreciable Risk Ratio <u>M-H, Fixed, 959</u> 1.14 [0.77, 1 1.06 [0.94, 1 1.18 [1.06, 1 1.16 [0.62, 2	harm or 69] .18] .31] .19]	no difference Risk Ra M-H, Fixed,	atio 95% Cl 		
• At a 95% cc <u>Study or Sub</u> 6.1.1 PDD Aarsland 200 Dubois 2012 Emre 2004 Ravina 2005 Subtotal (95%	onfidence bgroup 02 2 5 5% CI)	e level, data Cholinester Events 12 283 303 11	are consisten ase inhibitor 7 Total 14 3 377 3 362 21 774	t with ap Placel Events 9 123 127 9	bo Total 12 173 179 20 384	Weight 2.0% 35.1% 35.3% 1.9% 74.3 %	fit, appreciable Risk Ratio M-H, Fixed, 95% 1.14 [0.77, 1 1.06 [0.94, 1 1.18 [1.06, 1 1.16 [0.62, 2 1.12 [1.04, 1)	ham or 69] .18] .31] .19] 21]	no difference Risk Ra M-H, Fixed,	atio 95% Cl 		
• At a 95% co Study or Sult 6.1.1 PDD Aarsland 200 Dubois 2012 Emre 2004 Ravina 2005 Subtotal (95° Total events	onfidence bgroup 02 2 5 % CI)	e level, data Cholinester Events 12 283 303 11 609	are consisten ase inhibitor Total 14 3 377 3 362 21 774	t with ap Placel Events 9 123 127 9 268	bo Total 12 173 179 20 384	Weight 2.0% 35.1% 35.3% 1.9% 74.3%	fit, appreciable Risk Ratio <u>M-H, Fixed, 95%</u> 1.14 [0.77, 1 1.06 [0.94, 1 1.18 [1.06, 1 1.16 [0.62, 2 1.12 [1.04, 1]	harm or 6 CI .69] .18] .31] .19] 21]	no difference Risk Ra M-H, Fixed,	atio 95% Cl 		

Test for overall effect: Z = 2.97 (P = 0.003)

6.'	1.2	DLI	В	

lkeda 2015	64	96	31	46	8.7%	0.99 [0.77, 1.26]	
McKeith 2000	54	59	46	61	9.4%	1.21 [1.03, 1.43]	
Mori 2012 Subtotal (95% Cl)	83	105 260	24	34 141	7.5% 25.7 %	1.12 [0.88, 1.42] 1.11 [0.98, 1.25]	
Total events Heterogeneity: Chi² = 2.0 Test for overall effect: Z =	201 0, df = 2 (P = 0.37) 1.67 (P = 0.10)	; I² = 0%	101				
Total (95% CI)		1034		525	100.0%	1.12 [1.05, 1.19]	

 Total events
 810
 369

 Heterogeneity: Chi² = 4.00, df = 6 (P = 0.68); l² = 0%
 Test for overall effect: Z = 3.41 (P = 0.0007)
 Test for subgroup differences: Chi² = 0.01, df = 1 (P = 0.90), l² = 0%



20

PDD/DLB – cholinesterase inhibitor vs placebo: any adverse events (proportion of participants experiencing ≥1) – forest plot

	Cholinesterase inh	ibitor	Place	bo		Risk Ratio	Risk Ratio
Study or Subgroup	Events	Total	Events	Total	Weight	M-H, Fixed, 95% Cl	M-H, Fixed, 95% Cl
6.2.1 PDD							
Dubois 2012	67	377	22	173	36.5%	1.40 [0.89, 2.18]	+ -
Emre 2004	47	362	26	179	42.1%	0.89 [0.57, 1.39]	
Subtotal (95% CI)		739		352	78.6%	1.13 [0.82, 1.54]	◆
Total events	114		48				
Heterogeneity: Chi ² =	: 1.94, df = 1 (P = 0.16); I ^z = 48	1%				
Test for overall effect:	Z = 0.75 (P = 0.45)						
6.2.2 DLB							
lkeda 2015	5	96	5	46	8.2%	0.48 [0.15, 1.57]	
McKeith 2000	10	59	8	61	9.5%	1.29 [0.55, 3.05]	
Mori 2012	8	105	2	34	3.7%	1.30 [0.29, 5.81]	
Subtotal (95% CI)		260		141	21.4%	0.98 [0.53, 1.82]	
Total events	23		15				
Heterogeneity: Chi ² =	: 1.92, df = 2 (P = 0.38); I ^z = 09	6				
Test for overall effect:	Z = 0.06 (P = 0.95)						
Total (95% CI)		999		493	100.0%	1.10 [0.83, 1.45]	•
Total events	137		63				
Heterogeneity: Chi ^z =	4.00, df = 4 (P = 0.41); I ^z = 09	6				
Test for overall effect:	Z = 0.65 (P = 0.52)						U.US U.Z 1 5 ZU
Test for subgroup diff	ferences: Chi ² = 0.15,	df = 1 (i	^o = 0.69),	l ² = 0%	6		ravours medication ravours placebo

PDD/DLB – cholinesterase inhibitor vs placebo: serious adverse events (proportion of participants experiencing ≥1) – forest plot

	Cholinesterase inh	nibitor	Place	bo		Risk Ratio	Risk Ratio
Study or Subgroup	Events	Total	Events	Total	Weight	M-H, Fixed, 95% Cl	M-H, Fixed, 95% Cl
6.3.1 PDD							
Aarsland 2002	2	14	0	12	0.8%	4.33 [0.23, 82.31]	
Dubois 2012	58	377	19	173	40.1%	1.40 [0.86, 2.28]	+
Emre 2004	62	362	14	179	28.8%	2.19 [1.26, 3.80]	
Subtotal (95% CI)		753		364	69.7%	1.76 [1.23, 2.53]	
Total events	122		33				
Heterogeneity: Chi ² =	1.81, df = 2 (P = 0.40)); I ^z = 09	б (
Test for overall effect:	Z = 3.08 (P = 0.002)						
6.3.2 DLB							
lkeda 2015	11	96	5	46	10.4%	1.05 [0.39, 2.86]	_
McKeith 2000	7	59	7	61	10.6%	1.03 [0.39, 2.77]	
Mori 2012	7	105	4	34	9.3%	0.57 [0.18, 1.82]	
Subtotal (95% CI)		260		141	30.3%	0.90 [0.49, 1.63]	
Total events	25		16				
Heterogeneity: Chi ² =	0.78, df = 2 (P = 0.68	3); I z = 09	б				
Test for overall effect:	Z = 0.36 (P = 0.72)						
Total (95% CI)		1013		505	100.0%	1.50 [1.10, 2.04]	◆
Total events	147		49				
Heterogeneity: Chi ² =	6.09, df = 5 (P = 0.30)); I² = 18	3%				
Test for overall effect:	Z = 2.60 (P = 0.009)						U.05 U.2 1 5 20
Test for subgroup dif	ferences: Chi ² = 3.61	, df = 1 (P = 0.06),	l² = 72	.3%		Favours medication Favours placebo

PDD/DLB – cholinesterase inhibitor vs placebo: adverse events requiring treatment withdrawal (proportion of participants experiencing) – forest plot

PDD/DLB - cholinesterase inhibitor vs. placebo: cognitive outcomes

		Quali	ty assessment			No of	f patients	Effect	
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Chl	Placebo	Mean difference (95% Cl)	Quanty
MMSE – cholineste	rase inhibi	tors (follow-up 10	to 24 weeks; range	of scores: 0-30; hi	gher is better); se	e Figure	30 for forest	: plot	
7 ^{1–7}	RCT	not serious	not serious	not serious	not serious	1008	503	1.46 higher (1.11 to 1.82 higher)	⊕⊕⊕⊕ HIGH
MMSE – donepezil	(follow-up	10 to 24 weeks; ra	ange of scores: 0-30;	higher is better)					
51,2,4,6,7	RCT	not serious	not serious	not serious	not serious	614	276	1.68 higher (1.24 to 2.11 higher)	⊕⊕⊕⊕ HIGH
MMSE – rivastigmi	ne (follow-u	up 20 to 24 weeks	; range of scores: 0-	30; higher is bette	r)				
2 ^{3,5}	RCT	not serious	not serious	not serious	not serious	394	227	1.04 higher (0.43 to 1.65 higher)	⊕⊕⊕⊕ HIGH
¹ Aarsland 2002									

² Dubois 2012; data for 2 active treatment groups were combined (donepezil 5mg and 10mg). Mean and standard deviation calculated from data reported in paper

³ Emre 2004

⁴ Ikeda 2015; data for 2 active treatment groups were combined (donepezil 5mg and 10mg)

⁵ McKeith 2000

⁶ Mori 2012; data for 3 active treatment groups were combined (donepezil 3mg, 5mg and 10mg)

⁷ Ravina 2005

	Cholines	sterase inh	ibitor	PI	acebo			Mean Difference	Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Fixed, 95% Cl	IV, Fixed, 95% CI
6.4.1 PDD									
Aarsland 2002	22.8	3.7	12	21	5	12	1.0%	1.80 [-1.72, 5.32]	
Dubois 2012	22.974	3.009	386	21.427	3.009	170	42.9%	1.55 [1.00, 2.09]	-
Emre 2004	0.8	3.8	335	-0.2	3.5	166	28.1%	1.00 [0.33, 1.67]	
Ravina 2005 Subtotal (95% CI)	24.5	3.2	19 752	22.5	3.7	19 367	2.6% 74.7 %	2.00 [-0.20, 4.20] 1.36 [0.95, 1.77]	•
Heterogeneity: Chi² = Test for overall effect	: 1.95, df = 3 : Z = 6.48 (P	8 (P = 0.58) 4 < 0.00001	; I² = 0%)						
6.4.2 DLB									
lkeda 2015	1.817	3.1577	94	0.6	3	44	10.6%	1.22 [0.12, 2.31]	
McKeith 2000	0.67	4.26	59	-0.57	4.26	61	5.4%	1.24 [-0.28, 2.76]	+
Mori 2012 Subtotal (95% CI)	2.299	3.4988	103 256	-0.4	2.7	31 136	9.3% 25.3 %	2.70 [1.53, 3.87] 1.77 [1.06, 2.47]	 ◆
Heterogeneity: Chi ² =	: 3.89, df = 2	? (P = 0.14)	; I² = 49%						
restion overall ellect	. Z = 4.80 (F	~ 0.00001	,						
Total (95% CI)			1008			503	100.0%	1.46 [1.11, 1.82]	•
Heterogeneity: Chi ² = Test for overall effect	: 6.78, df = 6 : 7 = 8 06 (P) (P = 0.34) < 0 00001	; I² = 12%)						-10 -5 0 5 10
Test for subgroup dif	Terences: C	hi ² = 0.95.	, df=1(P=	= 0.33), l ^a	= 0%				Favours placebo Favours medication

PDD/DLB – cholinesterase inhibitor vs placebo: MMSE – forest plot

		Quali	ty assessment			No of p	patients	Effect (95% CI)	Quality			
No of studies	Design	Risk of bias	Inconsistency	cy Indirectness Imprecision		Chl	Placebo		Quanty			
Global function -	- cholines	terase inhibitor	s (follow-up 10 to	24 weeks; measi	ured with: CIBIC	+, ADCS-C	CGIC or CG	IC; range of scores: 1-7; lower is better); see Figure 31	for forest plot			
5 ^{1–5}	RCT	not serious	serious ⁶	not serious	not serious	798	396	SMD 0.48 lower (0.76 to 0.21 lower)	⊕⊕⊕O MODERATE			
Global function -	- donepez	il (follow-up 10	to 24 weeks; meas	ured with: CIBIC	C+, ADCS-CGIC	or CGIC; r	ange of sco	pres: 1-7; lower is better)				
4 1,2,3,5	RCT	not serious	serious ⁶	not serious	not serious	469	231	SMD 0.6 lower (1.08 to 0.11 lower)	⊕⊕⊕O MODERATE			
Global response plot	– choline	sterase inhibito	ors (at least minima	al improvement;	follow-up 10 to 3	24 weeks;	measured	with: CIBIC+ or ADCS-CGIC; higher is better); see Figur	e 32 for forest			
4 ^{1–4}	RCT	not serious	not serious	not serious	not serious	356/779 (45.7%)	129/377 (34.2%)	RR 1.31 (1.12 to 1.54) 106 more per 1000 (from 41 more to 185 more)	⊕⊕⊕⊕ HIGH			
Global response	– donepe	zil (at least min	imal improvement	; follow-up 10 to	24 weeks; meas	ured with:	CIBIC+ or	ADCS-CGIC; higher is better)				
31,2,4	RCT	not serious	serious ⁶	not serious	not serious	222/450 (49.3%)	80/212 (37.7%)	RR 1.27 (1.04 to 1.55) 102 more per 1000 (from 15 more to 208 more)	⊕⊕⊕O MODERATE			
¹ Aarsland 2002 ² Dubois 2012; ³ Emre 2004 ⁴ Mori 2012; da	2 data for 2 ta for 3 ad	active treatme ctive treatment	ent groups were co groups were com	ombined (donep bined (donepez	bezil 5mg and 1 til 3mg, 5mg an	0mg). Mei d 10mg)	an and sta	ndard deviation calculated from data reported in pap	er			

PDD/DLB - cholinesterase inhibitor vs. placebo: global assessment

⁵ Ravina 2005

⁶ Heterogeneity >40% between studies

	Cholinesterase inhibitor Placebo Std. Mean Differenc					Std. Mean Difference	Std. Mean Difference		
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Random, 95% Cl	IV, Random, 95% Cl
6.5.1 PDD									
Aarsland 2002	3.3	0.9	12	4.1	0.8	12	8.1%	-0.91 [-1.76, -0.06]	
Dubois 2012	3.65	1.2	347	3.9	1.27	170	30.6%	-0.20 [-0.39, -0.02]	-
Emre 2004	3.8	1.4	329	4.3	1.5	165	30.4%	-0.35 [-0.54, -0.16]	+
Ravina 2005 Subtotal (95% CI)	3.58	0.77	19 707	3.95	0.85	19 366	12.0% 81.1 %	-0.45 [-1.09, 0.20] - 0.30 [-0.45, -0.15]	•
Test for overall effect:	Z = 4.00 (P	< 0.0001)	5 (1 - 0.	55),1 =	13.0				
Mori 2012 Subtotal (95% CI) Heterogeneity: Not ap Test for overall effect:	3.0997 oplicable : Z = 4.54 (P	1.1512 < 0.00001)	91 91	4.27	1.197	30 30	18.9% 18.9 %	-1.00 [-1.43, -0.57] - 1.00 [-1.43, -0.57]	•
Total (95% CI) Heterogeneity: Tau ² = Test for overall effect:	= 0.06; Chi²: : Z = 3.45 (P	= 12.87, df: = 0.0006)	798 = 4 (P = (0.01); I²:	= 69%	396	100.0%	-0.48 [-0.76, -0.21]	-4 -2 0 2 4 Favours medication Favours placebo
Test for subgroup dif	ferences: C	hi² = 8.96, c	lf = 1 (P =	= 0.003)	. I ^z = 88	.8%			area an

PDD/DLB – cholinesterase inhibitor vs placebo: global function (different measures) – forest plot

(Cholinesterase inhibitor			bo		Risk Ratio Risk Ratio			
Study or Subgroup	Events	Total	Events	Total	Weight	M-H, Fixed, 95% Cl	M-H, Fixed, 95% Cl		
6.6.1 PDD									
Aarsland 2002	5	12	2	12	1.2%	2.50 [0.60, 10.46]			
Dubois 2012	155	347	68	170	52.6%	1.12 [0.90, 1.39]			
Emre 2004 Subtotal (95% CI)	134	329 688	49	165 347	37.6% 91.3 %	1.37 [1.05, 1.79] 1.24 [1.05, 1.47]	_ _		
Total events Heterogeneity: Chi² = 2. Test for overall effect: Z	294 35, df = 2 (P = 0.3 = 2.49 (P = 0.01)	1); I² = 16	119 5%						
6.6.2 DLB									
Mori 2012 Subtotal (95% CI)	62	91 91	10	30 30	8.7% 8.7 %	2.04 [1.21, 3.46] 2.04 [1.21, 3.46]			
Total events Heterogeneity: Not appl Test for overall effect: Z	62 icable = 2.67 (P = 0.008))	10						
Total (95% CI)		779		377	100.0%	1.31 [1.12, 1.54]	◆		
Total events Heterogeneity: Chi² = 5. Test for overall effect: Z	356 70, df = 3 (P = 0.1 = 3.29 (P = 0.001)	3); I² = 47 0)	129 '%				0.05 0.2 1 5 2		
Test for subaroup differ	ences: Chi ² = 3.16	6. df = 1 (P = 0.08).	. I ² = 68	.4%		Favours placebo Favours medication		

PDD/DLB – cholinesterase inhibitor vs placebo: global response (at least minimal improvement) – forest plot

		Quali	ty assessment			No o	of patients	Effect	Quality
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Chl	Placebo	Mean difference (95% CI)	Quanty
NPI-10 item – chol	inesterase	inhibitors (follow	w-up 12 to 24 weeks	; range of scores:	0-120; lower is b	etter) ¹ ;	see Figure 3	3 for forest plot	
5 ²⁻⁶	RCT	not serious7	not serious	not serious	not serious	931	465	1.49 lower (2.69 to 0.29 lower)	⊕⊕⊕⊕ HIGH
NPI-10 item – done	epezil (foll	ow-up 12 to 24 w	eeks; range of score	es: 0-120; lower is	better) ¹				
32,4,6	RCT	not serious7	serious ⁸	not serious	serious ⁹	550	246	0.92 lower (2.54 lower to 0.69 higher)	⊕⊕OO LOW
NPI-10 item – rivas	stigmine (f	ollow-up 20 to 24	weeks; range of sc	ores: 0-120; lower	is better)				
2 ^{3,5}	RCT	not serious	not serious	not serious	not serious	381	219	2.2 lower (4 to 0.39 lower)	⊕⊕⊕⊕ HIGH
UPDRS III – donep	ezil (follov	v-up 24 weeks; lo	ower is better); see F	igure 34 for fores	t plot				
4 4,6,10,11	RCT	serious ¹²	not serious	not serious	not serious ¹³	228	109	0.71 lower (2.09 lower to 0.66 higher)	⊕⊕⊕O MODERATE

PDD/DLB – cholinesterase inhibitor vs. placebo: other non-cognitive outcomes

¹ SD not reported for this outcome in Ikeda 2015; calculated from SE reported in paper

² Dubois 2012; data for 2 active treatment groups were combined (donepezil 5mg and 10mg). Mean and standard deviation calculated from data reported in paper

³ Emre 2004

⁴ Ikeda 2015; data for 2 active treatment groups were combined (donepezil 5mg and 10mg)

⁵ McKeith 2000

⁶ Mori 2012; data for 3 active treatment groups were combined (donepezil 3mg, 5mg and 10mg)

⁷ Data for this outcome not reported in Aarsland 2002. This represents a very small proportion of the total participants in the analysis, therefore quality assessment not downgraded

⁸ Heterogeneity > 40% between studies

⁹ At a 95% confidence level, data are consistent with appreciable benefit, appreciable harm or no difference

¹⁰ Aarsland 2002

¹¹ Ravina 2005

¹²Data for outcome not reported in 3 large RCTs (Dubois 2012, Emre 2004 and McKeith 2000). Papers stated no significant difference between groups

¹³CI do not cross the MID between 3.25 (Horvath et al., 2015) and 5 points (Schrag et al., 2006)

	Cholinesterase inhibitor Placebo						Mean Difference		Mean D	ifference			
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Fixed, 95% Cl		IV, Fixe	d, 95% Cl		
6.8.1 PDD													
Dubois 2012	11.711	10.32	354	13.055	10.32	170	40.7%	-1.34 [-3.23, 0.54]			+		
Emre 2004	-2	10	334	0	10.4	166	39.6%	-2.00 [-3.91, -0.09]			-		
Subtotal (95% CI)			688			336	80.3%	-1.67 [-3.01, -0.32]		•	•		
Heterogeneity: Chi ² =	0.23, df = 1	(P = 0.63);	l² = 0%										
Test for overall effect:	Z= 2.43 (F	9 = 0.01)											
6.8.2 DLB													
lkeda 2015	-4.4468	9.618	94	-6.4	9.95	44	11.7%	1.95 [-1.57, 5.48]		-	+		
McKeith 2000	-5	16.2	47	-1.2	10.7	53	4.9%	-3.80 [-9.25, 1.65]			+-		
Mori 2012	-5.8088	15.3238	102	0.3	17.5	32	3.2%	-6.11 [-12.86, 0.64]			+		
Subtotal (95% CI)			243			129	19.7%	-0.77 [-3.48, 1.94]					
Heterogeneity: Chi ² =	5.88, df = 2	2 (P = 0.05);	I² = 66%										
Test for overall effect:	Z=0.56 (F	9 = 0.58)											
Total (95% CI)			931			465	100.0%	-1.49 [-2.69, -0.29]		•	•		
Heterogeneity: Chi ² =	6.45, df = 4	(P = 0.17);	I ² = 38%						H		<u> </u>	+	
Test for overall effect:	Z = 2.43 (F	P = 0.02)							-20	-10 Favours modication	U Favoure pla	iU Icebo	20
Test for subaroup diff	, ferences: C	hi² = 0.34, c	#f = 1 (P =	= 0.56), l ^a	= 0%					ravours medication	r avours pia	neno	

PDD/DLB – cholinesterase inhibitor vs placebo: NPI-10 item – forest plot

Cholinesterase inhibitor				Placebo Mean Differenc					Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Fixed, 95% Cl	IV, Fixed, 95% Cl
6.9.1 PDD									
Aarsland 2002	31.8	15.4	12	35.1	8.1	12	2.0%	-3.30 [-13.14, 6.54]	
Ravina 2005	40.3	13.6	21	40.5	13.7	20	2.7%	-0.20 [-8.56, 8.16]	
Subtotal (95% CI)			33			32	4.7%	-1.50 [-7.87, 4.87]	
Heterogeneity: Chi ² =	0.22, df = 1	(P = 0.64);	I²=0%						
Test for overall effect:	Z=0.46 (P	= 0.64)							
6.9.2 DLB									
lkeda 2015	-0.6281	6.2929	96	-0.9	6.1	46	40.3%	0.27 [-1.89, 2.44]	-
Mori 2012	-0.6667	6.5098	99	0.7	3.8	31	55.1%	-1.37 [-3.22, 0.49]	
Subtotal (95% CI)			195			77	95.3%	-0.67 [-2.08, 0.73]	•
Heterogeneity: Chi ² =	1.27, df = 1	(P = 0.26);	i r = 21%)					
Test for overall effect:	Z=0.94 (P	= 0.35)							
Total (95% CI)			228			109	100.0%	-0.71 [-2.09, 0.66]	•
Heterogeneity: Chi ² =	1.55, df = 3	(P = 0.67);	I²=0%					-	
Test for overall effect:	Z=1.02 (P	= 0.31)							-20 -10 0 10 20
Test for subaroup diff	, ferences: C	hi² = 0.06. o	df=1 (P =	= 0.80).	l ^z = 0%	6			Favours medication Favours placebo

PDD/DLB – cholinesterase inhibitor vs placebo: UPDRS III – forest plot

Mixed population (PDD or DLB) – memantine

PDD/DLB – memantine vs. placebo: adverse events

		Quality	y assessment			No of pa	tients		Effect	
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Memantine	Placebo	Relative (95% Cl)	Absolute (95% CI)	Quality
Any adverse e	vents (p	robability of e	experiencing ≥1	; follow-up 16	to 24 weeks;	lower is bet	ter); see l	Figure 35 for forest plot		
2 ^{1,2}	RCT	not serious	not serious	not serious	serious ³	52/107 (48.6%)	52/113 (46%)	RR 1.06 (0.8 to 1.41)	28 more per 1000 (from 92 fewer to 189 more)	⊕⊕⊕O MODERATE
Serious advers	se event	s (probability	of experiencing	g ≥1; follow-u	p 16 to 24 wee	eks; lower is	better); s	ee Figure 36 for forest	plot	
2 ^{1,2}	RCT	not serious	not serious	not serious	serious ³	15/107 (14%)	11/113 (9.7%)	RR 1.43 (0.69 to 2.97)	42 more per 1000 (from 30 fewer to 192 more)	⊕⊕⊕O MODERATE
Adverse event	s requir	ing treatment	withdrawal (pro	bability of exp	periencing; fo	llow-up 16 to	o 24 week	s; lower is better); see	Figure 37 for forest plot	
2 ^{2,4}	RCT	not serious	not serious	serious⁵	serious ³	18/130 (13.8%)	21/137 (15.3%)	RR 0.91 (0.51 to 1.63)	14 fewer per 1000 (from 75 fewer to 97 more)	⊕⊕OO LOW
¹ Emre 2010;	data rep	ported for tot	al population (F	PDD and DLB)					

² Leroi 2009; not clear if adverse event data reported at end of active treatment (16 weeks) or end of drug withdrawal phase (22 weeks) ³ At a 95% confidence level, data are consistent with appreciable harm, appreciable benefit or no difference

⁴ Aarsland 2009

⁵ Both studies included people who were also taking a cholinesterase inhibitor



PDD/DLB – memantine vs placebo: any adverse events (proportion of participants experiencing ≥1) – forest plot

Error! No text of specified style in document.

	Meman	tine	Place	bo		Risk Ratio	Risk Ratio
Study or Subgroup	Events	Total	Events	Total	Weight	M-H, Fixed, 95% Cl	M-H, Fixed, 95% Cl
7.2.1 PDD							
Leroi 2009	1	11	1	14	8.2%	1.27 [0.09, 18.14]	
Subtotal (95% CI)		11		14	8.2%	1.27 [0.09, 18.14]	
Total events	1		1				
Heterogeneity: Not ap	plicable						
Test for overall effect:	Z = 0.18 (P = 0.8	6)				
7.2.2 DLB							
Emre 2010	14	96	10	99	91.8%	1.44 [0.67, 3.09]	
Subtotal (95% CI)		96		99	91.8%	1.44 [0.67, 3.09]	
Total events	14		10				
Heterogeneity: Not ap	plicable						
Test for overall effect:	Z = 0.95 (P = 0.3	4)				
Total (05% CI)		107		113	100.0%	1 43 10 60 2 071	
Total (95% CI)	4.5	107		115	100.0%	1.45 [0.09, 2.97]	
i otal events	15		11				
Heterogeneity: Chi*=	U.U1, df =	1 (P = 1	0.93); F=	0%			0.05 0.2 1 5 20
Test for overall effect:	Z = 0.96 (P = 0.3	4)				Favours medication Favours placebo
Test for subgroup diff	erences:	Chi² = C).01, df=	1 (P = I	0.93), I² =	0%	

PDD/DLB – memantine vs placebo: serious adverse events (proportion of participants experiencing ≥1) – forest plot



PDD/DLB – memantine vs placebo: adverse events requiring treatment withdrawal (proportion of participants experiencing) – forest plot

$\Gamma D D D D D D = \Pi H H H H H H H H H H H H H H H H H H$	PDD/DLB -	memantine	vs.p	lacebo:	cognitive	outcomes
--	-----------	-----------	------	---------	-----------	----------

		Qualit	ty assessment			No of pat	tients	Effect	Quality			
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Memantine	Placebo	Mean difference (95% CI)	Quality			
MMSE (follow-up	16 to 24 we	eeks; range of so	ores: 0-30; higher i	s better); see Fig	ure 38 for forest p	olot						
21,2	RCT not serious not serious serious ³ serious ³ 40 47 1.56 higher (0.17 lower to 3.28 higher) $\oplus \oplus OO$ LOW											
¹ Aarsland 2009 ² Leroi 2009; data ³ Both studies ind ⁴ At a 95% confid	Aarsland 2009 Leroi 2009; data reported for end of drug treatment phase (16 weeks) Both studies included people who were also taking a cholinesterase inhibitor At a 95% confidence level, data are consistent with appreciable harm, appreciable benefit or no difference											



PDD/DLB – memantine vs placebo: MMSE – forest plot

$\Gamma D D D D D = \Pi \Theta \Pi \Theta \Pi \Theta \Pi \Theta \Theta$

		Qualit	y assessment			No of pat	ients	Effect	Quality
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Memantine	Placebo	Standardised mean difference (95% CI)	Quanty
Global function (fe	ollow-up 2	24 weeks; measu	red with: ADCS-CO	GIC or CGIC; rang	e of scores: 1-7	; lower is better); see Figu	re 39 for forest plot	
2 ^{1,2}	RCT	not serious	not serious	not serious	not serious	123	130	0.27 lower (0.51 to 0.02 lower)	⊕⊕⊕⊕ HIGH
¹ Aarsland 2009									

² Emre 2010; data reported for total population (PDD and DLB)

	Memantine Placebo				9	Std. Mean Difference	Std. Mean Difference		
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Fixed, 95% Cl	IV, Fixed, 95% CI
7.5.1 DLB									
Emre 2010	3.5	1.4567	93	3.8	1.4885	97	75.6%	-0.20 [-0.49, 0.08]	
Subtotal (95% CI)			93			97	75.6%	-0.20 [-0.49, 0.08]	•
Heterogeneity: Not a	pplicable	9							
Test for overall effect	: Z = 1.39	9 (P = 0.1	6)						
7.5.2 Mixed									
Aarsland 2009	3.5	1.5	30	4.2	1.5	33	24.4%	-0.46 [-0.96, 0.04]	
Subtotal (95% CI)			30			33	24.4%	0.46 [-0.96, 0.04]	•
Heterogeneity: Not a	pplicable	9							
Test for overall effect	: Z = 1.80) (P = 0.0	7)						
Total (95% CI)			123			130	100.0%	-0.27 [-0.51, -0.02]	◆
Heterogeneity: Chi ² =	0.77, df	= 1 (P = 1	0.38); P	² =0%					
Test for overall effect	: Z = 2.10) (P = 0.0	4)						-4 -2 U 2 4
Test for subgroup dif	ferences	S: Chi² = 0).77, df	= 1 (P =	= 0.38), I ²	= 0%			Favours medication Favours placebo

PDD/DLB – memantine vs placebo: global function (different measures) – forest plot

	neman	une	; vs. pia	icen	0. aci	ivilies (JI Uai	יוא ווא	ng					
			Qua	lity as	sessme	nt				No of	patients	Effect		Quality
No of studies	Design	Ri	isk of bias	In	consist	ency lı	ndirect	ness	Imprecision	Memantine	Placebo	Standardised mean difference (9	5% CI)	Quality
ADL (follow-up 2	4 weeks;	mea	sured with	: ADC	S-ADL	or DAD; h	igher is	s better	; see Figure 4	0 for forest p	lot			
2 ^{1,2}	RCT	not	serious	not s	serious	not	seriou	s s	serious ³	123	130	0.13 higher (0.12 lower to 0.38 hi	gher)	⊕⊕⊕O MODERATE
¹ Aarsland 2009 ² Emre 2010; data reported for total population (PDD and DLB) ³ At a 95% confidence level, data are consistent with appreciable harm, appreciable benefit or no difference														
		M	emantine			Placebo			Std. Mean Di	fference		Std. Mean Difference		
Study or Subgro	oup M	ean	SD	Total	Mean	SD	Total	Weigh	t IV, Fixe	ed, 95% Cl		IV, Fixed, 95% Cl		
7.6.1 DLB														
Emre 2010 Subtotal (95% C	:1)	0	13.1101	93 93	-1.1	13.3965	97 97	75.39 75.3 9	6 0.08 (-0 6 0.08 (-0).20, 0.37] . 20, 0.37]				
Heterogeneity: N	Not applic	able												
Test for overall e	effect: Z =	0.57	(P = 0.57)											
7.6.2 Mixed														
Aarsland 2009 Subtotal (95% C	:1)	-1	6.4	30 30	-2.5	4.6	33 33	24.79 24.7 9	6 0.27 [-0 6 0.27 [-0).23, 0.76] . 23, 0.76]		•		
Heterogeneity: N Test for overall e	vot applic effect: Z =	able 1.06	(P = 0.29)											
Total (95% CI)				123			130	100.09	6 0.13 [-0	.12, 0.38]		•		
Heterogeneity: C	Chi² = 0.4	0, df=	= 1 (P = 0.9	53); I ř :	= 0%					-	-4	+ $+$ $+$ $+$ $+$ $+$ $+$ $+$ $+$ $+$		
Test for overall e	effect: Z =	1.02	(P = 0.31)								Favou	rs placebo Favours medication		

PDD/DLB - memantine vs. placebo: activities of daily living

Test for subgroup differences: Chi² = 0.40, df = 1 (P = 0.53), l² = 0%

PDD/DLB – memantine vs placebo: activities of daily living (different measures) – forest plot

PDD/DLB – memantine vs. placebo: carer-reported outcomes

		Qualit	ty assessment			No of pat	tients	Effect	Quality	
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Memantine	Placebo	Mean difference (95% CI)	Quanty	
ZBI (follow-up 16	ZBI (follow-up 16 to 24 weeks; lower is better); see Figure 41 for forest plot									
2 ^{1,2}	RCT	not serious	not serious	not serious	serious ³	104	111	2.69 lower (5.99 lower to 0.6 higher)	⊕⊕⊕O MODERATE	
¹ Emre 2010; dat ² Leroi 2009; dat	Emre 2010; data reported for total population (PDD and DLB) Leroi 2009; data reported for end of drug treatment phase (16 weeks)									

³ At a 95% confidence level, data are consistent with appreciable harm, appreciable benefit or no difference



PDD/DLB – memantine vs placebo: ZBI – forest plot

PDD/DLB – memantine vs. placebo: other non-cognitive outcomes

		Qualit	y assessment			No of pa	tients	Effect (05% CI)	Quality
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	n Memantine Placebo		Effect (95% Cl)	Quanty
NPI (follow-up 16	eks; measured	with: NPI-10 item o	or NPI 12-item; lo	ower is better) ¹ ;	see Figure 42	for forest p	lot		
2 ^{2,3}	RCT	not serious	not serious	not serious	serious ⁴	122	130	SMD 0.16 lower (0.41 lower to 0.08 higher)	⊕⊕⊕O MODERATE
UPDRS III (follow	/-up 16 to	24 weeks; lowe	r is better); see Fig	gure 43 for fores	t plot				
2 ^{2,3}	RCT	not serious	not serious	not serious	not serious⁵	131	141	MD 0.28 higher (1.28 lower to 1.85 higher)	⊕⊕⊕⊕ HIGH
¹ Data from Lerc ² Aarsland 2009 ³ Emre 2010; da ⁴ At a 95% confi ⁵ CI do not cross	oi 2009 co ata report idence le s the MID	ould not be incl ed for total pop vel, data are co between 3 (Ho	luded in this analy pulation (PDD and posistent with app prvath et al., 2015	vsis due to incol DLB) reciable harm, a) and 5 points (nsistent outcom appreciable ber Schrag et al., 2	ne reporting nefit or no diffe 006)	erence		

	М	emantine		Placebo			1	Std. Mean Difference	Std. Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Fixed, 95% Cl	IV, Fixed, 95% Cl
7.8.1 DLB									
Emre 2010 Subtotal (95% CI)	-2.6	13.5957	93 93	0.4	14.3889	97 97	75.4% 75.4 %	-0.21 [-0.50, 0.07] - 0.21 [-0.50, 0.07]	•
Heterogeneity: Not ap	plicable	!							
Test for overall effect:	Z=1.47	' (P = 0.14))						
7.8.2 Mixed									
Aarsland 2009 Subtotal (95% Cl)	-1.5	10.8	29 29	-1.4	10.6	33 33	24.6% 24.6 %	-0.01 [-0.51, 0.49] - 0.01 [-0.51, 0.49]	↓
Heterogeneity: Not ap	plicable	!							
Test for overall effect:	Z = 0.04	(P = 0.97))						
Total (95% CI)			122			130	100.0%	-0.16 [-0.41, 0.08]	•
Heterogeneity: Chi ² =	0.48, df	= 1 (P = 0.	49); l² =	= 0%					
Test for overall effect:	Z = 1.29	P = 0.20)						-4 -2 U Z 4 Eavours medication Eavours placebo
Test for subgroup diff	erences	: Chi² = 0.4	48, df=	1 (P = 0	0.49), I ^z = 0)%			ravours medication ravours placebo

PDD/DLB – memantine vs placebo: NPI (different measures) – forest plot

Error! No text of specified style in document.

	Mean Difference	Mean Difference			Placebo	I		emantine	M	
	IV, Fixed, 95% Cl	IV, Fixed, 95% Cl	Weight	Total	SD	Mean	Total	SD	Mean	Study or Subgroup
										7.9.1 PDD
		2.40 [-4.84, 9.64]	4.7%	14	9.1	21.9	10	8.8	24.3	Leroi 2009
		2.40 [-4.84, 9.64]	4.7%	14			10			Subtotal (95% CI)
									oplicable	Heterogeneity: Not a
)	i (P = 0.52)	Z = 0.65	Test for overall effect
										70200
										7.9.2 DLB
	—	-0.10 [-3.03, 2.83]	28.6%	97	10.4195	1.3	93	10.1968	1.2	Emre 2010
	—	-0.10[-3.03, 2.83]	28.0%	97			93			Subtotal (95% CI)
									oplicable	Heterogeneity: Not a
)	(P = 0.95)	: Z = 0.07	l est for overall effect
										7.9.3 Mixed
	+	0.30 [-1.62, 2.22]	66.7%	30	4.3	0	28	3.1	0.3	Aarsland 2009
	◆	0.30 [-1.62, 2.22]	66.7%	30			28			Subtotal (95% CI)
									oplicable	Heterogeneity: Not a
)	(P = 0.76)	Z = 0.31	Test for overall effect
		0.28 [.1.28, 1.85]	100.0%	141			131			Total (95% CI)
	T T		1001070			- 004	00\·IZ-	- 2 /P - 0	-0.20 HF	Hotorogonoity: Chiž-
)	-20 -10 0 10 20					- 0 %	02),1 -	ー Z (r ー 0. : /ロー 0 7つ	0.38, un 7 – 0.26	Teet for overall effect
	Favours medication Favours placebo			104		2/0-0	, 20. df	r (F = 0.72) ∴ Chi≅ = 0.1	. <u>2</u> - 0.30 Foroncoc	Test for cubgroup dif
)	-20 -10 0 10 20 Favours medication Favours placebo	2.40 [-4.84, 9.64] -0.10 [-3.03, 2.83] -0.10 [-3.03, 2.83] 0.30 [-1.62, 2.22] 0.30 [-1.62, 2.22] 0.30 [-1.62, 2.22]	4.7% 28.6% 28.6% 66.7% 66.7% 100.0%	97 97 97 30 30 141	10.4195 4.3 1.82), I² = 0	1.3 0 = 0% 2 (P = (10 93 93 93 28 28 28 131 82); ² =	(P = 0.52) 10.1968 (P = 0.95) 3.1 (P = 0.76) = 2 (P = 0.52) (P = 0.72) $: Chi^{2} = 0.52$	oplicable : Z = 0.65 1.2 oplicable : Z = 0.07 0.3 oplicable : Z = 0.31 : C.39, df : Z = 0.35 ferences	Subtotal (95% CI) Heterogeneity: Not a Test for overall effect 7.9.2 DLB Emre 2010 Subtotal (95% CI) Heterogeneity: Not a Test for overall effect 7.9.3 Mixed Aarsland 2009 Subtotal (95% CI) Heterogeneity: Not a Test for overall effect Total (95% CI) Heterogeneity: Chi ² = Test for overall effect Test for overall effect

PDD/DLB – memantine vs placebo: UPDRS III – forest plot

Network meta-analyses

Any adverse events

Quality assessment								
Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Quality			
Adverse events								
9 Aarsland 2002, Dubois 2012, Ikeda 2015, Mori 2012, Ravina 2005, Emre 2004, McKeith 2000, Emre 2010, Leroi 2009	Not serious	Not serious	Not serious ¹	Not serious	High			
¹ Considered not serious as population, intervention	ons. comparator and outcomes	are as defined in protocol						

Serious adverse events

Quality assessment								
Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Quality			
Serious adverse events								
7 Dubois 2012, Ikeda 2015, Mori 2012, Emre 2004, McKeith 2000, Emre 2010, Leroi 2009	Not serious	Not serious	Not serious ¹	Not serious	High			
¹ Considered not serious as population, intervention	ons, comparator and outcomes	are as defined in protocol						

Adverse events requiring treatment withdrawal

Quality assessment								
Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Quality			
Adverse events requiring treatment withdra	wal							
8 Aarsland 2002, Dubois 2012, Ikeda 2015, Mori 2012, Emre 2004, McKeith 2000, Aarsland 2009, Emre 2010	Not serious	Not serious	Not serious ¹	Not serious	High			
¹ Considered not serious as population, intervention	ons, comparator and outcomes	are as defined in protocol						

MMSE

Quality assessment									
Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Quality				
Change in MMSE scores									
9 Aarsland 2002, Dubois 2012, Ikeda 2015, Mori 2012, Ravina 2005, Emre 2004, McKaith 2000, Aarsland 2009, Emre	Not serious	Not serious	Not serious ¹	Not serious	High				
2010									
Considered not serious as population, intervention	ons, comparator and outcomes	are as defined in protocol							

Clincial global function

Quality assessment								
Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Quality			
Change in clinical global function (various measures)								
7	Not serious	Serious ¹	Not serious ²	Not serious	Moderate			

Quality assessment					
Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Quality
Aarsland 2002, Dubois 2012, Mori 2012, Ravina 2005, Emre 2004, Aarsland 2009, Emre 2010					
¹ Considerable between study heterogeneity (i ² > ⁴ ² Considered not serious as population, interventi	40%) ons, comparator and outcomes	are as defined in protocol			

NPI

Inconsistency	Indirectness	Imprecision	Quality	
Change in NPI scores				
Not serious	Not serious ¹	Not serious	High	
	Not serious	Not serious Not serious ¹	Inconsistency Indirectness Imprecision Not serious Not serious ¹ Not serious	

'Considered not serious as population, interventions, comparator and outcomes are as defined in protocol

UPDRS III (motor subscale)

Quality assessment					
Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Quality
Change in UPDRS III (motor) scores					
7 Aarsland 2002, Ikeda 2015, Mori 2012, Ravina 2005, Aarsland 2009, Emre 2010, Leroi 2009	Serious ¹	Not serious	Not serious ²	Serious ³	Low
¹ Some studies do not report measure of variation ² Considered not serious as population, interventions, comparator and outcomes are as defined in protocol ³ Analysis could not differentiate between any clinically distinct options					

Network meta-analyses

Mixed population (PDD or DLB)

PDD/DLB – any adverse events – FE model



Lower values favour treatment

Differences between treatments – relative risk and 95% confidence interval

	Placebo	Memantine	Chl	Donepezil	Rivastigmine
Placebo	N/A				
Memantine	1.05 (0.80, 1.39)	N/A			
Chl	1.14 (1.07, 1.22)	1.08 (0.81, 1.44)	N/A		
Donepezil	1.06 (0.97, 1.16)	1.01 (0.75, 1.35)	N/A	N/A	
Rivastigmine	1.19 (1.09, 1.30)	1.13 (0.84, 1.51)	N/A	1.12 (0.99, 1.27)	N/A

Quantifying heterogeneity/inconsistency:

tau^2 < 0.0001; l^2 = 0%

Test of heterogeneity/inconsistency:

Q	d.f.	p.value

1.31 6 0.971

Network graph:


PDD/DLB – serious adverse events – FE model



Differences between treatments – relative risk and 95% confidence interval

	Placebo	Memantine	Chl	Donepezil	Rivastigmine
Placebo	N/A				
Memantine	1.43 (0.69, 2.97)	N/A			
Chl	1.09 (0.82, 1.44)	0.76 (0.35, 1.67)	N/A		
Donepezil	1.23 (0.82, 1.84)	0.86 (0.37, 1.98)	N/A	N/A	
Rivastigmine	0.97 (0.65, 1.43)	0.68 (0.29, 1.55)	N/A	0.79 (0.45, 1.38)	N/A

Quantifying heterogeneity/inconsistency:

tau^2 < 0.0001; l^2 = 0%

Test of heterogeneity/inconsistency:

- Q d.f. p.value
- 3.3 4 0.5087



PDD/DLB – adverse events requiring treatment withdrawal – FE model



Lower values favour treatment

Differences between treatments – relative risk and 95% confidence interval

	Placebo	Memantine	Chl	Donepezil	Rivastigmine
Placebo	N/A				
Memantine	0.91 (0.51, 1.62)	N/A			
Chl	1.45 (1.06, 1.97)	1.59 (0.82, 3.05)	N/A		
Donepezil	1.22 (0.82, 1.84)	1.34 (0.66, 2.72)	N/A	N/A	
Rivastigmine	1.83 (1.13, 2.96)	2.01 (0.95, 4.26)	N/A	1.50 (0.80, 2.80)	N/A

Quantifying heterogeneity/inconsistency:

tau^2 < 0.0001; l^2 = 0%

Test of heterogeneity/inconsistency:

Q d.f. p.value

4.49 5 0.4819



PDD/DLB – MMSE – FE model



Higher values favour treatment

Differences between treatments – mean difference and 95% confidence interval

	Placebo	Memantine	Chl	Donepezil	Rivastigmine
Placebo	N/A				
Memantine	1.56 (-0.17, 3.28)	N/A			
Chl	1.46 (1.11, 1.82)	-0.09 (-1.85, 1.66)	N/A		
Donepezil	1.68 (1.24, 2.11)	0.12 (-1.66, 1.90)	N/A	N/A	
Rivastigmine	1.04 (0.43, 1.65)	-0.52 (-2.35, 1.31)	N/A	-0.64 (-1.39, 0.11)	N/A

Quantifying heterogeneity/inconsistency:

tau^2 < 0.0001; l^2 = 0%

Test of heterogeneity/inconsistency:

- Q d.f. p.value
- 5.15 6 0.5243



PDD/DLB – global function – RE model



Differences between treatments – standardised mean difference and 95% confidence interval

	Placebo	Memantine	Chl	Donepezil	Rivastigmine
Placebo	N/A				
Memantine	-0.31 (-0.78, 0.16)	N/A			
Chl	-0.50 (-0.81, -0.19)	-0.19 (-0.76, 0.37)	N/A		
Donepezil	-0.56 (-0.93, -0.20)	-0.25 (-0.85, 0.34)	N/A	N/A	
Rivastigmine	-0.35 (-0.92, 0.23)	-0.04 (-0.78, 0.70)	N/A	0.21 (-0.47, 0.90)	N/A

Quantifying heterogeneity/inconsistency:

tau^2 = 0.1182; l^2 = 70.7%

Test of heterogeneity/inconsistency:

Q d.f. p.value

13.63 4 0.0086



PDD/DLB – NPI – FE model



Lower values favour treatment

Differences between	treatments	 standardised mean 	difference an	nd 95% confidence	interval
----------------------------	------------	---------------------------------------	---------------	-------------------	----------

	Placebo	Memantine	Chl	Donepezil	Rivastigmine
Placebo	N/A				
Memantine	-0.16 (-0.41, 0.08)	N/A			
Chl	-0.15 (-0.27, -0.04)	0.01 (-0.26, 0.28)	N/A		
Donepezil	-0.11 (-0.26, 0.04)	0.06 (-0.23, 0.35)	N/A	N/A	
Rivastigmine	-0.21 (-0.38, -0.04)	-0.05 (-0.35, 0.25)	N/A	-0.10 (-0.33, 0.12)	N/A

Quantifying heterogeneity/inconsistency:

tau² = 0.0090; l² = 24.7%

Test of heterogeneity/inconsistency:

- Q d.f. p.value
- 5.31 4 0.2565



PDD/DLB – UPDRS III – FE model



Lower values favour treatment

Differences between treatments – mean difference and 95% confidence interval

	Placebo	Memantine	Chl	Donepezil
Placebo	N/A			
Memantine	0.28 (-1.28, 1.85)	N/A		
Chl	-0.71 (-2.09, 0.66)	-1.00 (-3.08, 1.09)	N/A	
Donepezil	-0.71 (-2.09, 0.66)	-1.00 (-3.08, 1.09)	N/A	N/A

Quantifying heterogeneity/inconsistency:

tau^2 < 0.0001; l^2 = 0%

Test of heterogeneity/inconsistency:

Q d.f. p.value

1.95 5 0.8566



PDD/DLB – UPDRS III sensitivity analysis – FE model

For this sensitivity analysis, in the 3 studies where the UPDRS III was measured but reported only as "non-significant", an effect size of 0 was assumed, and a SD imputed based on the pooled SD from the other trials of cholinesterase inhibitors versus placebo.



Lower values favour treatment

Differences between treatments – mean difference and 95% confidence interval

	Placebo	Memantine	ChI	Donepezil	Rivastigmine
Placebo	N/A				
Memantine	0.28 (-1.28, 1.85)	N/A			
Chl	-0.21 (-0.95, 0.53)	-0.49 (-2.22, 1.24)	N/A		
Donepezil	-0.34 (-1.29, 0.61)	-0.63 (-2.46, 1.21)	N/A	N/A	
Rivastigmine	0.00 (-1.18, 1.18)	-0.28 (-2.24, 1.68)	N/A	0.34 (-1.17, 1.86)	N/A

Quantifying heterogeneity/inconsistency:

tau^2 < 0.0001; l^2 = 0%

Test of heterogeneity/inconsistency:

Q d.f. p.value

2.48 7 0.9284



E.5 Non-pharmacological management of motor and non-motor symptoms

E.5.1 Physiotherapy and physical activity

Gait Outcomes

Outcome	No. of studies	Risk of bias	Inconsistency	Indirectness	Imprecision	Estimate (CI)	Overall quality
2 or 6 Minute Walk Test	10	Serious ¹	Not serious	Not serious	Not serious	MD 0.33 [0.11, 0.55]	Moderate
10 or 20m Walk Test	6	Serious ¹	Serious ²	Not serious	Serious ⁴	MD 0.02 [-0.63, 0.67]	Very Low
Speed	24	Serious ¹	Not serious	Not serious	Not serious	MD 0.06 [0.04, 0.08]	Moderate
Cadence (steps/min)	9	Serious ¹	Not serious	Not serious	Serious ⁴	MD 0.06 [-1.67, 1.78]	Low
Stride Length (m)	10	Serious ¹	Not serious	Not serious	Not serious	MD 0.06 [0.02, 0.10]	Moderate
Step Length (m)	7	Serious ¹	Not serious	Not serious	Serious ⁴	MD 0.02 [-0.00, 0.04]	Low
Freezing of Gait Questionnaire	4	Serious ¹	Not serious	Serious ³	Not serious	MD -1.41 [-2.63, -0.19]	Low

¹Individual study(ies) at risk of bias; ²Considerable between study heterogeneity (i²>40%); ³Serious indirectness: The GDG did not feel that the freezing of gate questionnaire was an adequate measure to quantify the severity and frequency of freezing in people with PD; ⁴Non-significant result

Functional Mobility and Balance Outcomes

Outcome	No. of studies	Risk of bias	Inconsistency	Indirectness	Imprecision	Estimate (CI)	Overall quality
Timed Up and Go	17	Serious ¹	Serious ²	Not serious	Serious ⁴	MD -1.09 [-1.57, -0.60]	Very Low
Functional Reach (cm)	6	Serious ¹	Serious ²	Not serious	Not serious	MD 2.82 [1.08, 4.55]	Low

Outcome	No. of studies	Risk of bias	Inconsistency	Indirectness	Imprecision	Estimate (CI)	Overall quality
Berg Balance Scale	11	Serious ¹	Serious ²	Not serious	Serious ⁵	MD 3.28 [1.96, 4.59]	Very Low
Activity Specific Balance Confidence	3	Serious ¹	Not serious	Not serious	Serious ⁶	MD 2.40 [-2.78, 7.57]	Low
Falls Efficacy Scale	8	Serious ¹	Serious ²	Serious ⁷	Serious ⁶	MD -3.59 [-7.55, 0.38]	Very Low
Number of people falling	2	Serious ¹	Serious ²	Not serious	Serious ⁶	OR 0.53 [0.20, 1.43]	Very Low

¹Individual study(ies) at risk of bias; ²Considerable between study heterogeneity ($i^2>40\%$); ³Serious indirectness: The GDG did not feel that the freezing of gate questionnaire was an adequate measure to quantify the severity and frequency of freezing in people with PD; ⁴Serious imprecision: MCIC = 11s was deemed clinically meaningful by the GDG; ⁵Serious imprecision: MCIC = 5 points was deemed clinically meaningful by the GDG; ⁶Non-significant results; ⁷Serious indirection: The GDG did not feel that the falls efficacy scale was an adequate measure to quantify the severity and frequency of falls in people with PD

Clinical-Rated Disability

Outcome	No. of studies	Risk of bias	Inconsistency	Indirectness	Imprecision	Estimate (CI)	Overall quality
UPDRS Total	7	Serious ¹	Serious ²	Not serious	Serious ³	MD -5.32 [-8.34, -2.30]	Very low
UPDRS Mental	4	Serious ¹	Not serious	Not serious	Not serious	MD -0.43 [-0.82, -0.05]	Moderate
UPDRS II (ADL)	7	Serious ¹	Not serious	Not serious	Not serious ⁴	MD -1.63 [-2.42, -0.84]	Moderate
UPDRS III (motor)	23	Serious ¹	Serious ²	Not serious	Serious⁵	MD -4.24 [-5.90, -2.58]	Very low

¹Individual study(ies) at risk of bias; ²Considerable between study heterogeneity (i²>40%); ³Cl cross the MID of 7.3 points (Schrag et al., 2006); ⁴Cl do not cross the MID of 3 points (Schrag et al., 2006); ⁵Cl cross the MID of 3.25 (Horvath et al., 2015) and 5 points (Schrag et al, 2006)

Clinical-rated QoL

Outcome	No. of studies	Risk of bias	Inconsistency	Indirectness	Imprecision	Estimate (CI)	Overall quality			
PDQ-39 Summary Index	14	Serious ¹	Serious ²	Not serious	Serious ⁴	MD -4.74 [-8.08, -1.39]	Very low			
PDQ-39 Mobility	4	Serious ¹	Not serious	Not serious	Serious ³	MD -2.31 [-6.55, 1.92]	Low			
¹ Individual study(ies) at risk of bias; ² Considerable between study heterogeneity (i ² >40%); ³ Non-significant result; ⁴ CI cross the MID of 1.6 points (Peto et al., 2001)										

PD REHAB (Clarke et al., 2016)

Outcome	No. of studies	Risk of bias	Inconsistency	Indirectness	Imprecision	Estimate (CI)	Overall quality
NEADL Summary Index (at 3 months)	1	Not serious	N/A	Serious ¹	Serious ²	MD 0.5 [-0.7, 1.7]	Low
NEADL Summary Index (at 15 months)	1	Not serious	N/A	Serious ¹	Serious ²	MD 0.07 [-0.64, 0.77]	Low
PDQ-39 Summary Index (at 3 months)	1	Not serious	N/A	Serious ¹	Not serious ³	MD 0.007 [-1.5, 1.5]	Moderate
PDQ-39 Summary Index (at 15 months)	1	Not serious	N/A	Serious ¹	Serious ⁴	MD -1.55 [-2.62, -0.47]	Low
EQ-5D quotient (at 3 months)	1	Not serious	N/A	Serious ¹	Serious ²	MD -0.03 [-0.07, -0.002]	Low
EQ-5D quotient (at 15 months)	1	Not serious	N/A	Serious ¹	Not serious	MD 0.02 [0.00007, 0.03]	Moderate
SF-12 physical (carers – at 3 months)	1	Not serious	N/A	Serious ¹	Serious ²	MD -0.6 [-2.3, 1.2]	Low
SF-12 mental (carers – at 3 months)	1	Not serious	N/A	Serious ¹	Not serious	MD -2.1 [-3.9, -0.3]	Moderate

Error! No text of specified style in document.

Outcome	No. of studies	Risk of bias	Inconsistency	Indirectness	Imprecision	Estimate (CI)	Overall quality				
¹ Considered serious as intervention is not as defined in protocol ² Non-significant result ³ CI does not cross the MID of 1.6 points (Peto et al., 2001) ⁴ CI cross											
the MID of 1.6 points (Peto et al., 2001)											

Alexander Technique

Outcome	No. of studies	Risk of bias	Inconsistency	Indirectness	Imprecision	Estimate (CI)	Overall quality
Self-assessment PD disability scale (SPDDS) at best	1	Not serious	N/A	Not serious	Not serious	MD -3.5 (-7.7 to - 0.0)	High
Self-assessment PD disability scale (SPDDS) at worst	1	Not serious	N/A	Not serious	Not serious	MD -6.3 (-11.8 to - 0.9)	High
Beck Depression Inventory	1	Not serious	N/A	Not serious	Serious ¹	MD -0.9 (-2.6 to 0.9)	Moderate
¹ Non-significant result	S						

Forest plots

Gait Outcomes

2 or 6 Minute Walk Test

	Inte	rvention No Intervention Std. Mean Difference		Std. Mean Difference	Std. Mean Difference				
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Fixed, 95% Cl	IV, Fixed, 95% Cl
1.1.1 Exercise v Con	trol								
Frazzitta 2015	411	105	16	405	93	15	9.8%	0.06 [-0.65, 0.76]	
Meek 2010	6.1	32.5	19	0.7	38.7	18	11.7%	0.15 [-0.50, 0.79]	
Schilling 2008	49.2	76.6	8	25.1	75.6	7	4.7%	0.30 [-0.72, 1.32]	
Schenkman 1998	15.4	35.7	23	1.9	42.5	23	14.4%	0.34 [-0.24, 0.92]	
Subtotal (95% CI)			66			63	40.6%	0.21 [-0.14, 0.56]	◆
Heterogeneity: Chi² =	: 0.43, df	= 3 (P =	0.93);	I ^z = 0%					
Test for overall effect	: Z = 1.19) (P = 0.2	23)						
1.1.2 Treadmill v Cor	ntrol								
Canning 2012	26.1	36	8	48.4	41.5	9	5.1%	-0.54 [-1.52, 0.43]	
Canning 2008	13.3	28.8	9	18.1	39.3	9	5.7%	-0.13 [-1.06, 0.79]	
Subtotal (95% CI)			17			18	10.8%	-0.33 [-1.00, 0.34]	-
Heterogeneity: Chi ² =	: 0.36, df	= 1 (P =	0.55);	I² = 0%					
Test for overall effect	: Z = 0.95	5 (P = 0.3	34)						
4.4.2 Dance u Contro	4								
T.T.J Dance V Contro Duman 204.0	"	400.0		24.0	07		40.40	0 00 / 0 04 0 701	
Duncan 2012	2.2	102.9	20	-21.6	97	20	10.4%	0.23 [-0.31, 0.78]	
Hackney 2009 Subtotal (95% CI)	54.Z	80.2	57	-7.5	127	43	13.3%	0.61 [0.01, 1.22] 0.40 [-0.00, 0.84]	
Hotorogonoity: Chiž -	- n op Af	- 1 /P -	10.00	IZ – ∩04		43	23.6 /4	0.40 [-0.00, 0.0 1]	
Tect for overall effect	· 0.03, ui · 7 – 1 06	- I (F - 10 - 0 C	0.30), 16)	1 - 0 %					
reaction over an effect	. 2 - 1.00	/ (i = 0.c	,5,						
1.1.4 Martial Arts v C	Control								
Hackney 2009	44.4	65.9	13	0.8	43.4	13	7.6%	0.76 [-0.04, 1.56]	
Choi 2013	472.1	58.6	11	368.6	152.6	9	5.6%	0.90 [-0.04, 1.83]	
Subtotal (95% CI)			24			22	13.2%	0.82 [0.21, 1.42]	
Heterogeneity: Chi ² =	: 0.05, df	= 1 (P =	0.83);	I ² = 0%					
l est for overall effect	: Z = 2.63	3 (P = 0.0	JU9)						
1.1.5 Nordic Walking	j v Contro	ol							
Cugusi 2015	395.6	78.4	10	329.2	54.9	10	5.6%	0.94 [0.00, 1.87]	
Subtotal (95% CI)			10			10	5.6%	0.94 [0.00, 1.87]	
Heterogeneity: Not a	pplicable								
Test for overall effect	: Z = 1.97	' (P = 0.0)5)						
Total (95% Cl)			174			156	100.0%	0.33 (0.11, 0.55)	•
Heterogeneity: Chiž -	:10.01 d	if = 10 (F	2 = 0.4	4): I≧ = 0	196				~
Test for overall effect	7 = 2 93	8/P=00	0.4						-2 -1 0 1 2
Test for subgroup dif	Terences	:Chi²=≀	8.35, d	lf = 4 (P	= 0.08),	I ² = 52	.1%		Favours No Intervention Favours Intervention

10 or 20m Walk test

	Inte	erventio	n	No In	itervent	ion	9	Std. Mean Difference	Std. Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Random, 95% Cl	IV, Random, 95% Cl
1.2.1 General Physic	therapy	v Contr	ol						
Stozek 2015 Subtotal (95% CI)	8.43	1.55	30 30	12.67	4.85	31 31	17.5% 17.5 %	-1.15 [-1.70, -0.61] - 1.15 [-1.70, -0.61]	•
Heterogeneity: Not a	pplicable								
Test for overall effect	: Z= 4.15	5 (P ≤ 0.	0001)						
1.2.2 Exercise v Con	trol								
Meek 2010	0	2.2	20	-0.7	3.8	18	16.7%	0.22 [-0.42, 0.86]	
Ni 2016a	0.13	0.166	27	-0.03	0.056	10	15.6%	1.07 [0.30, 1.84]	_
Schenkman 1998	0.1	0.2	23	-0.4	1	23	17.1%	0.68 [0.09, 1.28]	_
Stozek 2003	-1.3	1.8	30	0.2	4.9	31	17.8%	-0.40 [-0.91, 0.11]	
Subtotal (95% CI)			100			82	67.2%	0.36 [-0.28, 1.00]	
Heterogeneity: Tau² = Test for overall effect	= 0.32; CI : Z = 1.11	hi ^z = 12 (P = 0.	.69, df: 27)	= 3 (P =	0.005);	l² = 76'	%		
1.2.3 Treadmill v Cor	ntrol								
Kurtais 2008	-2.5	5.2	12	-1.7	3.7	12	15.3%	-0.17 [-0.97, 0.63]	
Subtotal (95% CI)			12			12	15.3%	-0.17 [-0.97, 0.63]	
Heterogeneity: Not a Test for overall effect	pplicable : Z = 0.42	? (P = 0.	68)						
Total (95% CI)			142			125	100.0%	0.02 [-0.63. 0.67]	
Heterogeneity: Tau ²	= 0.551 CI	hi ≅ = 32	25 df:	= 5 (P <	0 0000	1): I ² = 1	R4%		
Test for overall effect	· 7 = 0.07	'(P=0	95) 95		0.0000	·//· =·	0470		-2 -1 0 1 2
Test for subgroup dif	.∠–0.07 Terences	∵ Chi²=	1311	Favours Intervention Favours No Intervention					
reactor aubyloup un	rerentes	. 011 -	19.11,	ui – 2 (i	- 0.00	- 1. N	04.170		

Speed

	Intervention No Intervent		terventi	ention Mean Differe			Mean Difference		
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Fixed, 95% Cl	IV, Fixed, 95% Cl
1.3.1 General Physio	therapy	v Contro	DI						
Chandler 1999	0.13	0.36	26	0.1	0.3	26	1.1%	0.03 [-0.15, 0.21]	
Conradisson 2015	1.28	0.203	4/	1.17	0.199	44	5.4%	0.11 [0.03, 0.19]	
Ellis 2005	0.16	0.22	32	0.01	0.21	33	3.3%	0.15 [0.05, 0.25]	
Fisher 2008 Subtotal (95% CI)	0.02	0.18	10 115	0.02	U.17	10 113	1.6% 11.4 %	0.00 [-0.15, 0.15] 0.10 [0.04, 0.16]	•
Heterogeneity: Chi ² =	3.14, df	= 3 (P =	0.37);	I ² = 5%					
Test for overall effect:	Z = 3.42	(P = 0.0)	JOO6)						
1.3.2 Exercise v Con	trol								
Allen 2010	0.02	0.27	21	0.02	0.29	24	1.4%	0.00 [-0.16, 0.16]	
Boehm 2011	0.01	0.24	50	-0.002	0.23	52	4.4%	0.01 [-0.08, 0.10]	
Liao 2015	0.117	0.113	24	-0.009	0.039	12	14.5%	0.13 [0.08, 0.18]	-
Mak 2008	0.02	0.08	19	0	0.06	14	16.0%	0.02 [-0.03, 0.07]	+
Sage 2009a	0.06	0.2	31	-0.004	0.22	15	2.1%	0.06 [-0.07, 0.20]	
Thaut 1996	0.07	0.18	11	-0.05	0.27	11	1.0%	0.12 -0.07, 0.311	
Subtotal (95% CI)			156			128	39.4%	0.06 [0.03, 0.09]	◆
Heterogeneity: Chi ² =	11.24, d	f= 5 (P	= 0.05); I ^z = 56°	%				
Test for overall effect:	Z = 4.00	(P < 0.0	0001)						
1.3.3 Treadmill v Con	trol	_			_		_		
Canning 2008	0.09	0.12	9	0.06	0.09	9	3.8%	0.03 [-0.07, 0.13]	
Canning 2012	0.14	0.13	8	0.11	0.09	9	3.2%	0.03 [-0.08, 0.14]	
Fisher 2008	0.06	0.2	10	0.02	0.17	10	1.4%	0.04 [-0.12, 0.20]	
Pohl 2003	1.44	0.18	8	1.32	0.18	9	1.2%	0.12 [-0.05, 0.29]	
Protas 2005	0.17	0.35	9	0.01	0.23	9	0.5%	0.16 [-0.11, 0.43]	
Yang 2010	1.04	0.4	16	0.8	0.31	17	0.6%	0.24 [-0.01, 0.49]	
Subtotal (95% CI) Heterogeneity: Chi ² =	3.77, df	= 5 (P =	60 0.58);	I² = 0%		63	10.7%	0.06 [0.00, 0.12]	•
Test for overall effect:	Z = 2.00	(P = 0.0	05)						
1.3.4 Cueing v Contro	bl								
Almeida 2012	0.06	0.19	28	0.01	0.27	14	1.5%	0.05 [-0.11, 0.21]	
de Bruin 2010a	0.03	0.22	11	-0.02	0.17	11	1.4%	0.05 [-0.11, 0.21]	
Haase 2011	-0.05	0.26	17	0.048	0.2	6	0.9%	-0.10 [-0.30, 0.10]	
Mak 2008	0.05	0.06	19	0	0.06	14	21.3%	0.05 [0.01, 0.09]	-
Nieuwboer 2007	0.08	0.16	76	0.02	0.23	77	9.3%	0.06 [-0.00, 0.12]	
Thaut 1996	0.16	0.22	15	-0.05	0.27	11	1.0%	0.21 [0.02, 0.40]	
Subtotal (95% CI)			166			133	35.3%	0.05 [0.02, 0.09]	◆
Heterogeneity: Chi ² =	4.71, df	= 5 (P =	0.45);	² = 0%					
l est for overall effect:	Z = 3.24	(P = 0.)	JU1)						
1.3.5 Dance v Contro	0.05			0.00	0.00	47	4.00	0.001.040.000	
Hackney 2009 Subtotal (95% CI)	0.05	0.2	31	0.02	0.38	17	1.0% 1.0%	0.03 [-0.16, 0.22]	
Heterogeneity: Not ar	nlicable								-
Test for overall effect:	Z = 0.30	(P = 0.3	76)						
1.3.6 Martial Arts v C	ontrol								
Hackney 2009 Subtotal (95% CI)	0.01	0.21	13 13	0.1	0.11	13 13	2.2% 2.2 %	-0.09 [-0.22, 0.04] - 0.09 [-0.22, 0.04]	•
Heterogeneity: Not ap Test for overall effect:	plicable Z = 1.37	(P = 0.1	17)						
Total (95% Cl)			541			467	100.0%	0.06 [0.04, 0.08]	•
Heterogeneity: Chiz-	30.15 4	f = 23 /4	P = 0.1	5): IF = 24	1%				· · · · · · · · · · · · · · · · · · ·
Test for overall effect:	7 = 6.07	n – ∠o(n '(P < ∩)	0.1 – 0.1 100041	57,1 - 24	1.70				-1 -0.5 0 0.5 1
Test for subgroup diff	erences	∵Chi²=	7 28 c	lf = 5 (P =	= 0.20)	P= 31	3%		Favours No Intervention Favours Intervention

Cadence (steps/min)

	Intervention		No Intervention			Mean Difference			Mean Difference	
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Fixed, 95% Cl		IV, Fixed, 95% Cl
1.4.1 General Physic	therapy	v Contro	ol							
Conradsson 2015	115	6.782	46	113	6.633	44	38.8%	2.00 [-0.77, 4.77]		+ -
Fisher 2008	-1.6	10.9	10	0.8	8.9	10	3.9%	-2.40 [-11.12, 6.32]		
Subtotal (95% CI)			56			54	42.7%	1.60 [-1.05, 4.24]		◆
Heterogeneity: Chi ² =	: 0.89, df	= 1 (P =	0.35);	² = 0%						
Test for overall effect:	: Z = 1.18	3 (P = 0.2	24)							
4.4.2 Europie	I									
1.4.2 Exercise V Com	troi		~ ~ ~	_						
Sage 2009a	1.1	9.2	31	0	8.7	15	10.0%	1.10[-4.37, 6.57]		
Finaut 1996 Subtotal (05% CI)	-0.4	8	11	8.1	12	11	4.1%	-8.50 [-17.02, 0.02]		
Subtotal (95% CI)	0 45 df	- 4 /0 -	942	17 - 74 0	,	20	14.178	- 1.70 [-0.30, 2.90]		
Test for everall effect:	3.45, ui 7 = 0.73	= 1 (P =) / P = 0 4	0.06), 173	1-= / 19	0					
restior overall ellect.	. Z = 0.72	2 (P = 0.4	1)							
1.4.3 Treadmill v Con	ntrol									
Fisher 2008	0.2	9.6	10	0.8	8.9	10	4.5%	-0.60 [-8.71, 7.51]		
Protas 2005	7.5	7.7	9	6.6	14.2	9	2.7%	0.90 [-9.65, 11.45]		
Yang 2010	113	23	16	102	11	17	1.9%	11.00 [-1.42, 23.42]		
Subtotal (95% CI)			35			36	9.1%	2.29 [-3.42, 8.00]		
Heterogeneity: Chi ² =	2.44, df	= 2 (P =	0.29);	l ² = 189	6					
Test for overall effect:	Z = 0.79	9 (P = 0.4	3)							
1.4.4 Cuping v Contro	nl									
do Pruio 2010o	יי יי	7.0	11	1	12.6	11	2.000	1 00 [7 74 0 74]		
Le bruin 2010a	2.05	1170	17	0.5	0.70	11	3.870	1.00 [-7.74, 3.74]		
Nieuwhoer 2007	-0.4	10.2	76	2.5	11 9	77	2.7%	-2 90 [-6 41 0 61]		_ _
Thaut 1996	89	12.6	15	81	11 96	11	3 3 96	0.80[-8.72, 10.32]		
Subtotal (95% CI)	0.0	12.0	119	0.1	11.00	105	34.1%	-1.74 [-4.70, 1.21]		-
Heterogeneity: Chi ² =	1.45. df	= 3 (P =	0.69):	I ² = 0%				. / .		-
Test for overall effect:	Z=1.18	6 (P = 0.2	25)							
			<i>,</i>							
Total (95% CI)			252			221	100.0 %	0.06 [-1.67, 1.78]		•
Heterogeneity: Chi ² =	12.11, c	f = 10 (P	² = 0.2	8); I ^z = 1	7%					
Test for overall effect:	Z = 0.07	r (P = 0.9	95)						-20	-10 0 10 20 Favours Intervention
Test for subgroup dif	ferences	: Chi ² = 3	3.88, d	lf = 3 (P	= 0.28),	. I² = 22	.6%			

Stride Length (m)



Step Length (m)

	Inte	Intervention No Intervention		ion		Mean Difference	Mean Difference		
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Fixed, 95% Cl	IV, Fixed, 95% Cl
1.6.1 General Physio	therapy	v Contro	ol						
Amano 2013	0.55	0.11	15	0.59	0.06	9	9.1%	-0.04 [-0.11, 0.03]	
Conradsson 2015	0.67	1.32	46	0.62	1.35	44	0.1%	0.05 [-0.50, 0.60]	
Fisher 2008 Subtotal (95% CI)	0.01	0.08	10 71	0.03	0.11	10 63	5.9% 15.2 %	-0.02 [-0.10, 0.06] - 0.03 [-0.08, 0.02]	•
Heterogeneity: Chi² = Test for overall effect:	0.22, df Z = 1.17	= 2 (P = (P = 0.2	0.90); 24)	I² = 0%					
1.6.2 Exercise v Con	trol								
Boehm 2011	0.008	0.1	50	0.004	0.1	52	28.0%	0.00 [-0.03, 0.04]	+
Sage 2009a Subtotal (95% Cl)	0.03	0.1	31 81	0.003	0.11	15 67	9.7% 37.7 %	0.03 [-0.04, 0.09] 0.01 [-0.02, 0.04]	
Heterogeneity: Chi² = Test for overall effect:	0.35, df Z = 0.58	= 1 (P = (P = 0.5	0.56); 56)	l² = 0%					
1.6.3 Treadmill v Cor	ntrol								
Fisher 2008 Subtotal (95% Cl)	0.04	0.09	10 10	0.03	0.11	10 10	5.4% 5.4%	0.01 [-0.08, 0.10] 0.01 [-0.08, 0.10]	
Heterogeneity: Not ap Test for overall effect	oplicable	(P = 0.8	321						
	2 - 0.22	() = 0.0	,						
1.6.4 Cueing v Contro	ol								
Almeida 2012	0.051	0.091	28	0.02	0.13	14	7.3%	0.03 [-0.04, 0.11]	+
Nieuwboer 2007 Subtotal (95% CI)	0.04	0.1	76 104	0	0.12	77 91	34.4% 4 1.7 %	0.04 [0.01, 0.07] 0.04 [0.01, 0.07]	→
Heterogeneity: Chi ² = Test for overall effect:	0.04, df Z = 2.37	= 1 (P = (P = 0.0	0.83); 02)	I ² = 0%					
Total (95% Cl) Heterogeneity: Chi ² = Test for overall effect:	: 5.76, df : Z = 1.49	= 7 (P = (P = 0.1	266 0.57); 14)	I² = 0%		231	100.0 %	0.02 [-0.00, 0.04]	-1 -0.5 0 0.5 1 Eavours No Intervention
Test for subaroup dif	ferences	: Chi ^z =	5.15. d	lf = 3 (P :	= 0.16)	$ ^2 = 4^2$.8%		

Freezing of Gait Questionnaire

	Inter	ventio	on	No Int	ervent	tion		Mean Difference	Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Fixed, 95% Cl	IV, Fixed, 95% Cl
1.7.1 Exercise v Con	trol								
Allen 2010	-1.3	5.5	21	1.1	6	24	13.2%	-2.40 [-5.76, 0.96]	
Subtotal (95% CI)			21			24	13.2%	-2.40 [-5.76, 0.96]	
Heterogeneity: Not ap	pplicable								
Test for overall effect:	:Z=1.40	(P = 0).16)						
4700									
1.7.2 Cueing v Contro	01								_
Nieuwboer 2007	-0.95	4.74	76	-0.08	5.09	77	61.4%	-0.87 [-2.43, 0.69]	
Suptotal (95% CI)			70				61.4%	-0.87 [-2.43, 0.69]	
Heterogeneity: Not ap	pplicable								
Test for overall effect	:Z=1.09	(P = 0).27)						
1.7.3 Dance v Contro	ol								
Duncan 2012	-0.5	5.1	26	1.9	5.4	26	18.3%	-2.40 [-5.26, 0.46]	e
Hackney 2009	-0.5	5	31	1.2	8.95	17	7.0%	-1.70 [-6.30, 2.90]	
Subtotal (95% CI)	0.0	Ū	57		0.00	43	25.3%	-2.21 [-4.63, 0.22]	
Heterogeneity: Chi ² =	: 0.06. df:	= 1 (P	= 0.80): $ \mathbf{r} = 0.9$	6				
Test for overall effect	: Z=1.78	(P = 0).07)						
Total (95% CI)			154			144	100.0%	-1.41 [-2.63, -0.19]	\bullet
Heterogeneity: Chi ² =	: 1.27, df:	= 3 (P	= 0.74); I² = 0%	6				
Test for overall effect:	: Z = 2.26	(P = 0)).02)						Favoure Intervention Eavoure No Intervention
Test for subgroup dif	Terences	: Chi ≇÷	= 1.21,	df = 2 (F)	P = 0.5	5), I ² = ()%		

Functional Mobility and Balance Outcomes

Timed Up and Go

	Inte	rvention	1	No In	terventi	ion		Mean Difference	Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Random, 95% Cl	IV, Random, 95% Cl
2.1.1 General Physiot	herapy \	/ Control	I						
Stozek 2015 Subtotal (95% CI)	0.87	0.24	30 30	1.89	1.65	31 31	9.6% 9.6 %	-1.02 [-1.61, -0.43] - 1.02 [-1.61, -0.43]	→
Heterogeneity: Not ap	plicable								
Test for overall effect:	Z = 3.40	(P = 0.0)	007)						
2.1.2 Exercise v Cont	rol								
Boehm 2011	-0.4	5.2	50	-0.2	3.9	52	4.5%	-0.20 [-1.99, 1.59]	
Frazzitta 2015	8.5	3	16	7.3	1.2	15	5.1%	1.20 [-0.39, 2.79]	
Goodwin 2009	0.13	11.09	61	-0.48	13.92	62	1.1%	0.61 [-3.83, 5.05]	
Hashimoto 2015	9.1	1.9	17	10.2	2.4	14	5.3%	-1.10 [-2.65, 0.45]	
Klassen 2007	-1.3	2.5	17	-0.2	1.85	6	4.2%	-1.10 [-3.00, 0.80]	
Liao 2015	-2.1	1.55	24	1	1.9	12	6.5%	-3.10 [-4.34, -1.86]	
Ni 2016a	-1.781	2.588	27	0.7	0.96	10	6.9%	-2.48 [-3.62, -1.34]	
Sage 2009a	-0.6	2.21	31	0	2.33	15	5.8%	-0.60 [-2.01, 0.81]	
Schilling 2008	-0.1	0.7	8	-0.75	1.2	7	7.6%	0.65 [-0.36, 1.66]	+
Stozek 2003 Subtotal (95% Cl)	-2.36	2.63	30 281	1.1	7.15	31 224	2.6% 49.5 %	-3.46 [-6.15, -0.77] -1.00 [-2.05, 0.06]	
Heterogeneity: Tau ² =	2.07: Ch	ni ^z = 39.8	1. df=	9 (P < (0.00001): $ ^2 = 7$	7%		
Test for overall effect:	Z = 1.85	(P = 0.0)	6)	•					
2.1.3 Cueing v Contro	d								
Almeida 2012	-1.45	3.2	28	0.1	3.16	14	3.8%	-1.55 [-3.59, 0.49]	
Nieuwboer 2007 Subtotal (95% Cl)	-1.59	4.59	76 104	-1.34	5.78	77 91	4.9% 8.7 %	-0.25 [-1.90, 1.40] - 0.77 [-2.05, 0.52]	-
Heterogeneity: Tau² = Test for overall effect:	0.00; Ch Z = 1.17	ni² = 0.94 (P = 0.24	4) 4)	(P = 0.	33); I² =	0%			
2.1.4 Dance v Control									
Hackney 2009	-1.1	4.31	31	2	9.28	17	1.0%	-3.10 [-7.76, 1.56]	
Hashimoto 2015	9.7	2.1	15	10.2	2.4	14	4.9%	-0.50 [-2.15, 1.15]	_
Subtotal (95% CI)			46			31	5.9%	-0.85 [-2.58, 0.89]	-
Heterogeneity: Tau ² =	0.19; Ch 7 – 0.96	i ^z = 1.06 (P = 0.3)	i, df = 1 4)	(P = 0.	30); I² =	6%			
rescior overall ellect.	2 - 0.30	(i = 0.5	4)						
2.1.5 Martial Arts v Co	ontrol			0.00		-	0.45		
Choi 2013	7.03	0.9	11	9.32	4.16	9	2.4%	-2.29 [-5.06, 0.48]	
Gao 2013	-1.38	1.2	37	0.03	1.15	39	9.9%	-1.41 [-1.94, -0.88]	-
Hackney 2009 Subtotal (95% CI)	-1	0.1	13 61	-0.1	1.1	13 61	9.6% 21.9%	-0.90 [-1.50, -0.30] - 1.21 [-1.63, -0.79]	↓
Heterogeneity: Tau² = Test for overall effect:	0.01; Ch Z = 5.62	ni² = 2.16 (P < 0.00	i, df = 2 0001)	? (P = 0.	34); I² =	7%			
2.1.6 Nordic Walking	v Contro	ol.							
Cuqusi 2015	8.1	1.9	10	10.1	2.3	10	4.3%	-2.00 (-3.85, -0.15)	
Subtotal (95% CI)			10		2.0	10	4.3%	-2.00 [-3.85, -0.15]	\bullet
Heterogeneity: Not ap Test for overall effect:	plicable Z = 2.12	(P = 0.03	3)					. ,	-
Total (95% Cl)			532			448	100.0%	-1.09 [-1.57, -0.60]	•
Heterogeneity Tau ² =	0.56: Ch	ni² = 45 9	 12 df=	18 (P =	0.0003): P= 6	1%		
Test for overall effect: Test for subgroup diff	Z = 4.35 erences:	(P < 0.0) Chi ² = 1	.2, ur – 001) .59. df	= 5 (P =	= 0.90).	/, - = 0 ² = 0%			-10 -5 Ó Ś 10 Favours Intervention Favours No Intervention

Functional Reach (cm)

	Intervention No Intervention				ion		Mean Difference	Mean Difference	
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Random, 95% Cl	IV, Random, 95% Cl
2.2.1 Exercise v Con	trol								
Ashburn 2007	0.4	6.56	67	-1	7	66	21.0%	1.40 [-0.91, 3.71]	- -
Ni 2016a	2.711	5.478	27	0.1	2.501	10	19.2%	2.61 [0.03, 5.19]	
Schenkman 1998	1.57	4.45	23	-0.28	4.17	23	19.8%	1.85 [-0.64, 4.34]	+ - -
Stozek 2003	5.8	6.01	30	-0.03	7.38	31	14.8%	5.83 [2.46, 9.20]	
Subtotal (95% CI)			147			130	74.9%	2.63 [0.95, 4.32]	◆
Heterogeneity: Tau ² =	= 1.14; C	hi = 4.8	39, df =	3 (P = 0).18); I ^z =	= 39%			
Test for overall effect:	: Z = 3.08	6 (P = 0.	002)						
2.2.2 Cueing v Contro	ol								
Nieuwboer 2007	1.8	5.28	76	0.34	9.02	77	20.8%	1.46 [-0.88, 3.80]	
Subtotal (95% CI)			76			- 77	20.8%	1.46 [-0.88, 3.80]	-
Heterogeneity: Not ap	pplicable	•							
Test for overall effect:	: Z = 1.22	2 (P = 0.	22)						
2.2.3 Nordic Walking	ı v Contr	ol							
Cuqusi 2015	-2.8	7.8	10	-14.2	9,9	10	4.3%	11.40 (3.59, 19.21)	
Subtotal (95% CI)			10			10	4.3%	11.40 [3.59, 19.21]	
Heterogeneity: Not as	oplicable								
Test for overall effect:	Z = 2.86	6 (P = 0.	004)						
			· ·						
Total (95% CI)			233			217	100.0%	2.82 [1.08, 4.55]	◆
Heterogeneity: Tau ² =	= 2.34; C	hi ² = 10	.62, df=	= 5 (P =	0.06); P	²= 53%	,		
Test for overall effect:	Z = 3.19	9 (P = 0.	001)	-					-20 -10 0 10 20 Economic No Intervention Economic Intervention
Test for subgroup dif	ferences	: Chi ^z =	5.77, d						

Berg Balance Scale

Error! No text of specified style in document.

	Inte	erventior	1	No In	tervent	ion		Mean Difference	Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Random, 95% Cl	IV, Random, 95% Cl
2.3.1 General Physic	otherapy	v Contro	d i						
Landers 2016 Subtotal (95% CI)	47.8	8.8	31 31	47.6	4.9	10 10	5.9% 5.9 %	0.20 [-4.14, 4.54] 0.20 [-4.14, 4.54]	
Heterogeneity: Not a	pplicable								
Test for overall effect	t: Z = 0.09	9 (P = 0.9	3)						
2.3.2 Exercise v Con	ntrol								
Ashburn 2007	1.5	9.51	67	1.6	10.21	66	8.1%	-0.10 [-3.45, 3.25]	
Goodwin 2009	3.1	11.07	61	-0.55	9.88	62	7.2%	3.65 [-0.06, 7.36]	
Hashimoto 2015	49.7	4.7	17	51.6	4.5	14	8.3%	-1.90 [-5.15, 1.35]	
Ni 2016a	4.304	3.084	27	0.4	0.807	10	14.7%	3.90 [2.64, 5.17]	-
Qutubuddin 2013	48	10.4	13	47.9	7.2	10	2.8%	0.10 [-7.10, 7.30]	
Taheri 2011 Subtotal (95% Cl)	8.4	9.7	12 197	-1.08	3.7	12 174	3.9% 44.9 %	9.48 [3.61, 15.35] 2.33 [-0.42, 5.09]	→
Heterogeneity: Tau ² :	= 7.74; C	hi² = 19.4	40, df=	= 5 (P =	0.002);	l ² = 749	%		
Test for overall effect	t: Z = 1.60	6 (P = 0.1	0)						
2.3.3 Treadmill v Co	ntrol								
Cakit 2007 Subtotal (95% Cl)	7.09	8.5	21 21	-1.42	10.07	10 10	2.8% 2.8 %	8.51 [1.29, 15.73] 8.51 [1.29, 15.73]	
Heterogeneity: Not a	pplicable								
Test for overall effect	t: Z = 2.31	(P = 0.0	2)						
2.3.4 Dance v Contro	ol								
Hackney 2009	3.95	4.7	31	-1.2	9.32	17	5.3%	5.15 [0.42, 9.88]	
Hashimoto 2015 Subtotal (95% CI)	55.1	1.2	15 46	51.6	4.5	14 31	10.7% 16.0 %	3.50 [1.07, 5.93] 3.85 [1.68, 6.01]	•
Heterogeneity: Tau ^z : Test for overall effect	= 0.00; C t: Z = 3.48	hi² = 0.33 3 (P = 0.0	7, df = 005)	1 (P = 0	1.54); I² =	= 0%			
2.3.5 Martial Arts v (Control								
Gao 2013	4.16	3.83	37	0.38	2.5	39	14.1%	3.78 [2.32, 5.24]	
Hackney 2009 Subtotal (95% Cl)	3.3	3	13 50	-0.5	2.1	13 52	12.2% 26.3 %	3.80 [1.81, 5.79] 3.79 [2.61, 4.97]	•
Heterogeneity: Tau ² : Test for everall offer	= 0.00; C	hi ² = 0.00), df =	1 (P = 0	1.99); i ² =	= 0%			
restion overall ellect	ι. <u>Ζ</u> = 0.3ι) (F < 0.0	0001)						
2.3.6 Nordic Walking	g v Contr	ol							
Cugusi 2015 Subtotal (95% Cl)	50.8	5.02	10 10	42.2	7.8	10 10	4.0% 4.0%	8.60 [2.85, 14.35] 8.60 [2.85, 14.35]	
Heterogeneity: Not a Test for overall effect	pplicable t: Z = 2.93) 3 (P = 0.0	03)						
T-4-1 (05%) OB			0.55			207	400.07	0.0014.00 4.50	
Total (95% CI)	- 2 62. 0		355 15 df	- 40.00	- 0.002	287	100.0%	3.28 [1.96, 4.59]	
Test for overall effect	= 2.62; U h 7 = 4.00	nr=∠8.0)(P ≈ 0 0	10, at : 100043	= 12 (P :	= 0.005)	i, in = 57	70		-20 -10 0 10 20
Test for subaroup di	. ∠ – 4.30 fferences	:Chi²=1	7.98. d	lf = 5 (P	= 0.16).	I² = 37	.3%		Favours No Intervention Favours Intervention

Activity Specific Balance Confidence

	Intervention			No In	itervent	ion		Mean Difference	Mean Difference		
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Fixed, 95% Cl	IV, Fixed, 95% Cl		
2.4.1 Exercise v Con	ntrol										
Klassen 2007	1.75	6.76	17	-1.7	6.87	6	66.0%	3.45 [-2.92, 9.82]			
Schilling 2008	3.3	8.35	8	-1.1	15.85	7	15.6%	4.40 [-8.69, 17.49]			
Subtotal (95% CI)			25			13	81.6%	3.63 [-2.09, 9.36]			
Heterogeneity: Chi ² =	= 0.02, df	= 1 (P	= 0.90); I ² = 09	б						
Test for overall effect	: Z = 1.24	4 (P = 0	D.21)								
2.4.2 Cueing v Contr	ol										
Shankar 2008	-2.1	16.5	14	1	16.1	14	18.4%	-3.10 [-15.18, 8.98]			
Subtotal (95% Cl)			14			14	18.4%	-3.10 [-15.18, 8.98]			
Heterogeneity: Not a	pplicable	9									
Test for overall effect	: Z = 0.50) (P = (D.61)								
Total (95% Cl)			39			27	100.0%	2.40 [-2.78, 7.57]			
Heterogeneity: Chi ² =	= 0.99, df	= 2 (P	= 0.61); I ² = 09	%						
Test for overall effect	t: Z = 0.91	(P = (D.36)						-20 -10 U 10 20 Eavoure No Intervention Eavoure Intervention		
Test for subaroup dif	fferences	: Chi ² ∶	= 0.97.	df = 1.0	P = 0.32	2). $I^2 = 0$	196				

<u>Falls</u>

Falls Efficacy Scale

	Inte	erventior	ı	No In	tervent	tion		Mean Difference	Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Random, 95% Cl	IV, Random, 95% Cl
3.1.1 General Physic	otherapy	v Contro	l						
Conradsson 2015 Subtotal (95% Cl)	27.3	8.227	47 47	26.5	7.96	44 44	21.5% 21.5 %	0.80 [-2.53, 4.13] 0.80 [-2.53, 4.13]	★
Heterogeneity: Not a	pplicable	/D = 0.6	-						
restion overall ellect	L Z = 0.47	(F = 0.0	4)						
3.1.2 Exercise v Con	ntrol								
Allen 2010	-2.3	10.6	21	1.3	10.56	24	15.7%	-3.60 [-9.80, 2.60]	
Goodwin 2009	-0.83	9.82	61	1.12	9.98	63	21.2%	-1.95 [-5.44, 1.54]	
Liao 2015	-9.1	7.1	24	0.8	6.6	12	18.7%	-9.90 [-14.59, -5.21]	
Subtotal (95% CI)			106			99	55.6%	-5.09 [-10.24, 0.07]	◆
Heterogeneity: Tau ² :	= 14.79; C	≎hi² = 7.2	22, df =	: 2 (P = 0	0.03); I ^z	= 72%			
Test for overall effect	t: Z = 1.93	(P = 0.0	5)						
3 1 3 Troadmill y Co	etrol								
J. I.J Tredumini V Cu	40.07	20.00	- 24		00.70	4.0	0.400	440710044 0 771	
Cakit 2007 Subtotal (05% CI)	-12.27	39.06	21	2.4	28.78	10	2.4%	-14.67 [-39.11, 9.77]	
Subtotal (95% CI)			21			10	2.470	- 14.07 [-39.11, 9.77]	
Heterogeneity: Not a	ppiicapie	(D 0 0							
Test for overall effect	1. Z = 1.18	(P = 0.2	4)						
3.1.4 Cueing v Contr	ol								
Nieuwboer 2007	4.48	25.37	76	1.16	29.44	77	11.4%	3.32 [-5.38, 12.02]	
Subtotal (95% Cl)			76			- 77	11.4%	3.32 [-5.38, 12.02]	
Heterogeneity: Not a	pplicable								
Test for overall effect	t: Z = 0.75	(P = 0.4)	5)						
3 1 5 Martial Arts v (ontrol								
Nocera 2013	-5 9	15.7	16	10	<u>g o</u>	a	Q N %	-10 60 6-21 22 0 021	
Subtotal (95% Cl)	-3.0	13.7	15	4.0	0.0	6	9.0%	-10.60 [-21.22, 0.02]	
Hotorogeneity: Not a	nnlicable					0	0.070	- loido [-e liee; olde]	
Teet for overall effect	PPIICADIE F 7 – 1 06	/P = 0.0	6)						
reactor overall ellect	. 2 - 1.80	(i = 0.0	3)						
Total (95% CI)			265			236	100.0%	-3.59 [-7.55, 0.38]	•
Heterogeneity: Tau ² :	= 16.12; 0	Chi² = 18	.44, df	= 6 (P =	0.005)	; i² = 67	'%		
Test for overall effect	: Z = 1.77	(P = 0.0	8)						-50 -25 0 25 50 Execute Intervention Execute No Intervention
Test for subgroup dif	fferences	: Chi² = 8	3.79, di	f=4 (P=	= 0.07),	I ² = 54.	5%		Favours intervention Favours No Intervention

Number of people falling

	Interver	ntion	No Interve	ention		Odds Ratio	Odds Ratio
Study or Subgroup	Events	Total	Events	Total	Weight	M-H, Random, 95% Cl	M-H, Random, 95% Cl
3.2.1 Exercise v Com	trol						
Canning 2015 Subtotal (95% CI)	75	115 115	81	116 116	58.7% 58.7 %	0.81 [0.47, 1.41] 0.81 [0.47, 1.41]	
Total events	75		81				
Heterogeneity: Not ap	plicable						
Test for overall effect:	Z=0.75 (P = 0.4	5)				
3.2.2 Martial Arts v C	ontrol						
Gao 2013	8	37	19	39	41.3%	0.29 [0.11, 0.79]	
Subtotal (95% CI)		37		39	41.3%	0.29 [0.11, 0.79]	
Total events	8		19				
Heterogeneity: Not ap	plicable						
Test for overall effect:	Z= 2.42 (P = 0.02	2)				
Total (95% Cl)		152		155	100.0%	0.53 [0.20, 1.43]	
Total events	83		100				
Heterogeneity: Tau ² =	0.36; Chi	² = 3.09	, df = 1 (P =	0.08); P	²= 68%		
Test for overall effect:	Z=1.26 ($P = 0.2^{\circ}$	1)				U.UT I 10 100 Esvoure Intervention
Test for subgroup diff	erences:	Chi ^z = 3	.08, df = 1 (P = 0.08	l), I ² = 67.	6%	

Clinical-Rated Disability

UPDRS Total

Intervention No Intervention Mean Difference Mean Differen	ce
Study or Subgroup Mean SD Total Mean SD Total Weight IV, Random, 95% Cl IV, Random, 95	% CI
4.1.1 General Physiotherapy v Control	
Cholewa 2013 31.7 11.85 40 42.1 10.35 30 12.7% -10.40 [-15.62, -5.18]	
Ellis 2005 -6.2 6.2 32 -1 6 33 17.2% -5.20 [-8.17, -2.23]	
Fisher 2008 -5.2 8.72 10 -3.2 10.1 10 8.1% -2.00 [-10.27, 6.27]	
Subtotal (95% Cl) 82 73 38.0% -6.30 [-10.39, -2.21]	
Heterogeneity: Tau ² = 6.43; Chi ² = 3.91, df = 2 (P = 0.14); l ² = 49%	
Test for overall effect: Z = 3.02 (P = 0.003)	
4.1.2 Exercise v Control	
Boehm 2011 -7.1 13.5 50 5.8 14.3 52 12.4% -12.90 [-18.30, -7.50]	
Hashimoto 2015 33.9 12.3 17 37.3 11.9 14 7.8% -3.40 [-11.95, 5.15]	
Park 2014 1.31 6.29 16 -0.13 8.43 15 12.7% 1.44 [-3.82, 6.70]	
Subtotal (95% Cl) 83 81 32.8% -5.02 [-14.56, 4.52]	-
Heterogeneity: Tau ² = 60.15; Chi ² = 14.11, df = 2 (P = 0.0009); i ² = 86%	
Test for overall effect: Z = 1.03 (P = 0.30)	
4.1.3 Treadmill V Control	
Fisher 2008 -2.1 14 10 -3.2 10.1 10 5.7% 1.10 [-9.60, 11.80]	
Ganesan 2015 34.75 5.83 20 41.5 5.03 20 15.7% -6.75 [-10.47, -3.03]	
Subutal (95% CI) 50 50 21.4% -4.49 [-11.40, 2.47]	
Heterogeneity 1aur=14.11; Chir=1.85, dt=1 (P=0.17); P=46%	
Test for overall effect: $Z = 1.27$ (P = 0.21)	
4.14 Dance v Control	
Trastilliou 2015 53.9 12.3 17 37.3 11.9 14 7.676 -3.40 [11.95, 5.15]	
Subtration for the row steep rises and	
Testion over an energy 2 = 0.70 (7 = 0.44)	
Total (95% Cl) 212 198 100.0% -5.32 [-8.34, -2.30]	
Heterogeneity: Tau ² = 11 54: Chi ² = 20 32, df = 8 (P = 0.009): i ² = 61%	
Test for overall effect $Z = 345$ ($P = 0.0006$) $- 0.0007$ $- 0.$	10 20
Favours Intervention Favo	urs No Intervention

UPDRS Mental

	Intervention			No Int	ervent	ion		Mean Difference	Mean Difference					
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Fixed, 95% Cl	IV, Fixed, 95% Cl					
4.2.1 General Physic	otherapy	v Con	trol											
Cholewa 2013	2.15	1.31	40	2.6	1.27	30	40.1%	-0.45 [-1.06, 0.16]						
Ellis 2005	-1.1	1.6	32	-0.5	1.3	33	29.6%	-0.60 [-1.31, 0.11]						
Fisher 2008	0.1	1.37	10	0.3	0.91	10	14.3%	-0.20 [-1.22, 0.82]						
Subtotal (95% CI)			82			73	84.1%	-0.46 [-0.88, -0.04]						
Heterogeneity: Chi ² :	Heterogeneity: Chi ² = 0.40, df = 2 (P = 0.82); i ² = 0%													
Test for overall effect	t: Z = 2.14	4 (P =)	D.O3)											
4.2.2 Treadmill v Co	ntrol													
Fisher 2008	0	1.97	10	0.3	0.91	10	8.2%	-0.30 [-1.64, 1.04]	_					
Subtotal (95% CI)	-		10			10	8.2%	-0.30 [-1.64, 1.04]						
Heterogeneity: Not a	pplicable	9												
Test for overall effect	t: Z = 0.44	4 (P =)	D.66)											
4.2.3 Martial Arts v	Control													
Choi 2013	1.27	1.84	11	1.56	1.33	9	7.7%	-0.29 [-1.68, 1.10]						
Subtotal (95% CI)			11			9	7.7%	-0.29 [-1.68, 1.10]						
Heterogeneity: Not a	pplicable	9												
Test for overall effect	t: Z = 0.41	I (P = I	D.68)											
Total (95% CI)			103			92	100.0%	-0.43 [-0.82, -0.05]						
Heterogeneity: Chi ² :	= 0 49 df	= 4 (P	= 0.97	Y = 0				,						
Test for overall effect	t 7 = 2 20) (P = 1	0.01	// - 0 /					-2 -1 0 1 2					
Test for subaroun di	fferences	v (Chi²	= N N9	df = 2 (F	2 = N 94	5) P= (1%		Favours Intervention Favours No Intervention					

UPDRS ADL

Study or Subgroup Mean SD Total Mean SD Total Weight IV, Fixed, 95% Cl 4.3.1 General Physiotherapy v Control Control 0 10.1% -4.58 [-7.06, -2.10] - - Cholewa 2013 12.02 5.26 40 16.6 5.24 30 10.1% -4.58 [-7.06, -2.10] - Conradsson 2015 12.3 4.799 47 13.2 4.643 44 16.6% -0.90 [-2.84, 1.04] - - Ellis 2005 -2.1 2.8 32 -0.3 2.3 33 40.0% -1.80 [-3.05, -0.55] - - Fisher 2008 -1.5 2.81 10 -0.9 3.9 10 7.0% -0.60 [-3.58, 2.38] -
4.3.1 General Physiotherapy v Control Cholewa 2013 12.02 5.26 40 16.6 5.24 30 10.1% -4.58 [-7.06, -2.10] Conradsson 2015 12.3 4.799 47 13.2 4.643 44 16.6% -0.90 [-2.84, 1.04] Ellis 2005 -2.1 2.8 32 -0.3 2.3 33 40.0% -1.80 [-3.05, -0.55] Fisher 2008 -1.5 2.81 10 -0.9 3.9 10 7.0% -0.60 [-3.58, 2.38] Subtotal (95% Cl) 129 117 73.7% -1.86 [-2.78, -0.94] - Heterogeneity: Chi ² = 6.24, df = 3 (P = 0.10); I ² = 52% Test for overall effect: Z = 3.97 (P < 0.0001) - - -
Cholewa 2013 12.02 5.26 40 16.6 5.24 30 10.1% -4.58 [-7.06, -2.10] Conradsson 2015 12.3 4.799 47 13.2 4.643 44 16.6% -0.90 [-2.84, 1.04] Ellis 2005 -2.1 2.8 32 -0.3 2.3 33 40.0% -1.80 [-3.05, -0.55] Fisher 2008 -1.5 2.81 10 -0.9 3.9 10 7.0% -0.60 [-3.58, 2.38] Subtotal (95% Cl) 129 117 73.7% -1.86 [-2.78, -0.94] \bullet Heterogeneity: Chi ² = 6.24, df = 3 (P = 0.10); I ² = 52% Test for overall effect: Z = 3.97 (P < 0.0001) \bullet \bullet \bullet \bullet
Conradsson 2015 12.3 4.799 47 13.2 4.643 44 16.6% -0.90 [-2.84, 1.04] Ellis 2005 -2.1 2.8 32 -0.3 2.3 33 40.0% -1.80 [-3.05, -0.55] Fisher 2008 -1.5 2.81 10 -0.9 3.9 10 7.0% -0.60 [-3.58, 2.38] Subtotal (95% Cl) 129 117 73.7% -1.86 [-2.78, -0.94] \bullet Heterogeneity: Chi ² = 6.24, df = 3 (P = 0.10); I ² = 52% Test for overall effect: Z = 3.97 (P < 0.0001)
Ellis 2005 -2.1 2.8 32 -0.3 2.3 33 40.0% -1.80 [-3.05, -0.55] Fisher 2008 -1.5 2.81 10 -0.9 3.9 10 7.0% -0.60 [-3.58, 2.38] Subtotal (95% Cl) 129 117 73.7% -1.86 [-2.78, -0.94] • Heterogeneity: Chi ² = 6.24, df = 3 (P = 0.10); I ² = 52% Test for overall effect: Z = 3.97 (P < 0.0001)
Fisher 2008 -1.5 2.81 10 -0.9 3.9 10 7.0% -0.60 [-3.58, 2.38] Subtotal (95% Cl) 129 117 73.7% -1.86 [-2.78, -0.94] Heterogeneity: Chi² = 6.24, df = 3 (P = 0.10); I² = 52% Test for overall effect: Z = 3.97 (P < 0.0001)
Subtotal (95% Cl) 129 117 73.7% -1.86 [-2.78, -0.94] Heterogeneity: Chi ² = 6.24, df = 3 (P = 0.10); l ² = 52% Test for overall effect: Z = 3.97 (P < 0.0001)
Heterogeneity: Chi² = 6.24, df = 3 (P = 0.10); l² = 52% Test for overall effect: Z = 3.97 (P < 0.0001)
Test for overall effect: Z = 3.97 (P < 0.0001)
4.3.2 Exercise v Control
Frazzitta 2015 6 2.6 16 7 4 15 10.9% -1.00 [-3.39, 1.39]
Subtotal (95% Cl) 16 15 10.9% -1.00 [-3.39, 1.39]
Heterogeneity: Not applicable
Test for overall effect: Z = 0.82 (P = 0.41)
4.3.3 Treadmill v Control
Fisher 2008 0.6 3.64 10 -0.9 3.9 10 5.7% 1.50 [-1.81, 4.81]
Subtotal (95% Cl) 10 10 5.7% 1.50 [-1.81, 4.81]
Heterogeneity: Not applicable
Test for overall effect; Z = 0.89 (P = 0.37)
4.3.4 Dance v Control
Duncan 2012 -0.6 7.5 26 1.9 8.4 26 3.3% -2.50 [-6.83, 1.83]
Subtotal (95% Cl) 26 26 3.3% -2.50 [-6.83, 1.83]
Heterogeneity: Not applicable
Test for overall effect: Z = 1.13 (P = 0.26)
4.3.5 Martial Arts v Placebo
Choi 2013 5.82 3.37 11 8.22 3.7 9 6.4% -2.40 5.53 0.73
Subtotal (95% Cl) 11 9 6.4% -2.40 [-5.53, 0.73]
Heterogeneity: Not applicable
Test for overall effect: Z = 1.50 (P = 0.13)
Total (95% CI) 192 177 100.0% -1.63 L-2.42 -0.841 -
Heteronenity $Chi^2 = 10.50$ df = 7 ($P = 0.16$) ($P = 34\%$
Testing worall different $7 = 40.6 (P \le 0.001)$ -34.00 -10 -5 0 5 10
Test for subgroup differences: ChiE = 4.35. df = 4 (P = 0.36) IE = 8.0%

UPDRS Motor

	Inter	vention		No In	terventi	on		Mean Difference	Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Random, 95% Cl	IV, Random, 95% Cl
4.4.1 General Physiother	apy v Cor	itrol							
Chandler 1999	-1	7	26	3	6.24	26	4.8%	-4.00 [-7.60, -0.40]	
Cholewa 2013	17.52	6.45	40	22.9	6.22	30	5.2%	-5.38 [-8.37, -2.39]	
Ellis 2005	-3	6.6	32	-0.2	5.3	33	5.2%	-2.80 [-5.72, 0.12]	
Fisher 2008 Subtotal (05% CI)	-3.8	8.17	10	-2.7	8.15	10	2.9%	-1.10 [-8.25, 6.05]	
Subtotal (93% CI)	0.068-1	0.00 46	- 2 /0	- 0.695	17 - 0.07	99	10.178	-3.07 [-3.02, -2.11]	\bullet
Teet for everall effect: 7 =	4 22 /P ~	2.08, 011	= 3 (P : \	= 0.56),	1-= 0.%				
Testion overall ellect. Z -	4.33 (F N	0.0001,	,						
4.4.2 Exercise v Control									
Boehm 2011	-7	8.6	50	6	9.5	52	4.9%	-13.00 [-16.51, -9.49]	
Frazzitta 2015	8.4	3.7	16	10.6	5.2	15	5.1%	-2.20 [-5.40, 1.00]	
Ni 2016a	-10.796	5.07	27	0.4	2.904	10	5.4%	-11.20 [-13.82, -8.57]	
Park 2014	2.13	5.43	16	0	7.45	15	4.2%	2.13 [-2.48, 6.74]	_
Qutubuddin 2013	10.4	4.8	13	14.7	7.1	10	3.9%	-4.30 [-9.42, 0.82]	
Sage 2009a	-3.7	6.72	31	1.2	7.81	15	4.2%	-4.90 [-9.51, -0.29]	
Subtotal (95% CI)			153			117	27.6%	-5.73 [-10.40, -1.06]	
Heterogeneity: Tau ² = 29.	.91; Chi ² =	46.58	df = 5	(P < 0.0	0001); I ^z	= 89%	•		
Lest for overall effect: $Z =$	2.40 (P =	0.02)							
4.4.3 Treadmill v Control									
Conning 2000	24	5.1	0	2.2	4.0	0	4 206	0.101462.4721	
Canning 2008 Canning 2012	1.5	3.1	8	2.3	4.3	a	4.2.70	0.10 [4.52, 4.72]	
Fisher 2008	-2.8	972	10	-27	815	10	2.6%	-0.10[-7.96_7.76]	
Ganesan 2015	25.1	4.79	20	30.35	3.8	20	5.3%	-5.25 [-7.93, -2.57]	
Subtotal (95% CI)			47			48	16.5%	1.60 [-5.19, 2.00]	
Heterogeneity: Tau ² = 7.8	5; Chi² = 1	7.76, df	= 3 (P :	= 0.05);	I ² = 61%				
Test for overall effect: Z =	0.87 (P =	0.38)							
4.4.4 Cueing v Control									
Almeida 2012	1.4	8.57	28	2.1	9.28	14	3.5%	-0.70 [-6.51, 5.11]	
de Bruin 2010a	-5.6	9.17	11	-1.8	6.53	11	3.1%	-3.80 [-10.45, 2.85]	
Shankar 2008 Subtotal (05% Ch	-4	8.29	14	1.21	1.15	14	3.5%	-5.21 [-11.15, 0.73]	
Hotorogonoity: Tou ² - 0.0	0: Chiž – 1	1 1 0 df	- 270-	- 0.663	12 - 0.00	35	10.174	-3.13[-0.00, 0.37]	
Test for overall effect: 7 -	1 75 (P -	n nev	- 2 (F	- 0.55),	1 - 0 %				
	1.10 (1 -	0.00)							
4.4.5 Dance v Control									
Duncan 2012	-12.7	12.1	26	-3	9.4	26	3.5%	-9.70 [-15.59, -3.81]	
Hackney 2009	-2.1	10.96	31	5	10.33	17	3.3%	-7.10 [-13.34, -0.86]	
Subtotal (95% CI)			57			43	6.8%	-8.48 [-12.76, -4.19]	
Heterogeneity: Tau² = 0.0	0; Chi ² = (0.35, df	= 1 (P :	= 0.55);	I ² = 0%				
Test for overall effect: Z =	3.88 (P =	0.0001))						
4.4.6 Martial Arta u Contr	al								
4.4.0 Martial Arts V Conu	01 00 4	47	45	22	5.0		4.200	4 40 5 3 00 6 70	
Amanu zu i s Obai 2012	25.4	4.7	10	16 4 4	0.0	9	4.3%	1.40 [-2.96, 5.76]	
Geo 2013	-9.05	9.73	37	-1 0	9.00	20	2.470	-0.00 [-3.00, 7.40]	
Hackney 2009	-0.05	6.6	13	43	5.6	13	4.1%	-5.80 [-10.51 -1.09]	
Schmitz-Hubsch 2006	-0.32	10.9	31	5.54	14 77	21	2.8%	-5.86 [-13.25.1.53]	
Subtotal (95% CI)			107			91	19.2%	-3.70 [-7.11, -0.28]	-
Heterogeneity: Tau ² = 8.3	6; Chi ² = 1	10.01, d	f= 4 (F	P = 0.04); I ² = 60	%			
Test for overall effect: Z =	2.12 (P =	0.03)							
4.4.7 Nordic Walking v C	ontrol								
Cugusi 2015	18.8	12.3	10	26	11.9	10	1.8%	-7.20 [-17.81, 3.41]	
Suprotal (95% CI)			10			10	1.8%	-7.20 [-17.81, 3.41]	
Heterogeneity: Not applic	able	0.4.00							
i est for overall effect: Z =	1.33 (P =	U.18)							
Total (95% CD			535			447	100.0%	-4.24 [-5 90 -2 59]	•
Heterogeneity Tour - 11	55' Chiž-	85.62	df= 24	(P < ∩)	000013-	ירד = דו ורק = דו		-424 [-3:30, -2:30]	
Test for overall effect: 7 =	5.01 (P <	00.00, 0.0000	ur – 24 1)	- ve i = 0.1	00001),	- 72			-20 -10 0 10 20
Test for subaroun differen	nces: Chiž	= 7.05	df = 6	(P = 0.3	(2), $ \mathbf{r} = 1$	4.9%			Favours Intervention Favours No Intervention
. sector subgroup differen		1.00,	ai = 0		-0.1 - 1	1.0.70			

Clinical-Rated QoL

PDQ-39 Summary Index

	Inte	rventio	n i	No Ir	nterventi	on		Mean Difference	Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Random, 95% Cl	IV, Random, 95% Cl
5.1.1 General Physic	therapyv	/ Contro	I						
Chandler 1999	4	14.94	26	3.32	12.65	26	7.1%	0.68 [-6.84, 8.20]	_
Cholewa 2013	33.17	12.62	40	46.16	13.68	30	8.0%	-12.99 [-19.26, -6.72]	
Ni 2016b	-2.5	4.773	14	5.2	10.487	10	7.5%	-7.70 [-14.66, -0.74]	
Subtotal (95% CI)			80			66	22.6%	-6.89 [-14.60, 0.83]	
Heterogeneity: Tau² =	: 34.03; C	¦hi² = 7.4	19, df =	2 (P = 0	0.02); I ² =	73%			
Test for overall effect:	Z = 1.75	(P = 0.0)	8)						
5.1.2 Exercise v Con	trol								
Allen 2010	-1	14.3	21	4.9	26.7	24	4.4%	-5.90 [-18.21, 6.41]	
Klassen 2007	0.25	4.06	17	-1	5.54	6	9.0%	1.25 [-3.58, 6.08]	_ _
Liao 2015	-15.45	13.95	24	2	3.6	12	8.2%	-17.45 [-23.39, -11.51]	
Meek 2010	-2.6	15.6	19	-3.1	17.42	17	5.1%	0.50 [-10.35, 11.35]	
Qutubuddin 2013	30.6	20.8	13	48.7	31.9	10	1.8%	-18.10 [-40.88, 4.68]	
Suptotal (95% CI)			94			69	28.6%	-6.99 [-16.99, 3.02]	
Heterogeneity: Tau ² =	: 97.69; C	≎hi² = 25	.35, df 	= 4 (P <	0.0001)	; l² = 84	%		
Test for overall effect:	Z=1.37	(P = 0.1	7)						
5.4.3 Troadmill v Con	trol								
Osenine 2000	0.5			4.0	5.4		7.50	0 70 / 7 00 0 000	
Canning 2008 Conning 2013	0.5	9.4	9	1.2	5.1	9	1.5%	-0.70[-7.09, 0.29]	
Subtotal (95% CI)	-3.7	3.4	17	2.9	5.2	9	9.0%	-6.20 [-10.33, -2.07] -4 20 [-9 39 0 98]	
Hotorogonoity: Tou ² -	6 66' CH	uZ – 1 70	tr df− '	1 /P = 0	10\12	1006	10.0 /4	-4.20 [-5.55, 0.50]	
Test for overall effect:	7 = 1.59	(P = 0.1)	7, ur – 1)	r (r = 0.	10,1 - 1	+5 /0			
restion overall ellect.	2 = 1.55	(1 = 0.1	·/						
5.1.4 Cueing v Contro	ol								
- Nieuwboer 2007	-3.42	11.08	76	-1.84	13.28	77	9.6%	-1.58 [-5.45, 2.29]	+ _
Subtotal (95% CI)			76			77	9.6%	-1.58 [-5.45, 2.29]	-
Heterogeneity: Not ap	plicable								
Test for overall effect:	Z = 0.80	(P = 0.4)	2)						
5.1.5 Dance v Contro									
Hackney 2009	-3.84	5.4	31	-1.5	13.05	17	7.8%	-2.34 [-8.83, 4.15]	
Subtotal (95% CI)			31			17	7.8%	-2.34 [-8.83, 4.15]	
Heterogeneity: Not ap	oplicable								
Test for overall effect:	Z = 0.71	(P = 0.4	8)						
E.4.6 Mortial Arta u.C	ontrol								
5.1.0 Martial Arts V C	UNUU		4.0		40.05		7.00		
Hackney 2009	1.55	5.37	13	-1.5	13.05	17	7.6%	3.05 [-3.81, 9.91]	
Nocera 2013 Subtotal (95% CN	1.1	7.5	15 29	y	8.5	5 23	5.9% 14.5%	-7.90 [-15.69, -0.11] 2.26 [12.00, 8.46]	
Hotorogonaity Tav2-	45.04.0	hii - Ar	20)0 df-	1/0 - (- 51 - 74 - 7	∠J 770/.	14.070	-2.20 [-12.99, 0.40]	
Tect for everall effect:	- 40.94; U 7 = 0.44	/D=0.0	:o,ur= o\	- (⊢ = l	5.04), 11=	1170			
restior overall ellect.	∠= 0.41	(r ² = 0.6	0)						
Total (95% CI)			326			270	100.0%	-4.74 [-8.081.39]	•
Heteroneneity: Tau? =	: 26 27 [.] C	:hi² = 45	38 df	= 13 (P	< 0.0001): P= 7	1%		
Test for overall effect:	7=2.78	(P = 0.0		100	0.0001	//· - /			-20 -10 0 10 20
Test for subaroun diff	ferences:	Chi ≧ = 0.0	2.34. di	f = 5 (P =	= 0.80) I ^a	'= 0%			Favours Intervention Favours No Intervention
PDQ-39 Mobility

	Inte	erventio	n	No In	tervent	ion		Mean Difference	Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Fixed, 95% Cl	IV, Fixed, 95% Cl
5.2.1 General Physic	therapy	v Contr	ol						
Keus 2007b	4.11	14.5	14	-2.12	12.2	13	17.7%	6.23 [-3.85, 16.31]	
Ni 2016b	-2.4	9.259	14	1.1	6.534	10	45.0%	-3.50 [-9.82, 2.82]	
Subtotal (95% CI)			28			23	62.6%	-0.76 [-6.11, 4.60]	
Heterogeneity: Chi ² =	2.57, df	= 1 (P =	0.11);	l² = 619	6				
Test for overall effect:	Z = 0.28	(P = 0.1	78)						
5005 · 0									
5.2.2 Exercise v Com	trol								
Subtotal (95% CI)			U			U		Not estimable	
Heterogeneity: Not ap	oplicable								
l est for overall effect:	: Not app	licable							
5.2.3 Dance v Contro)l								
Hackney 2009	-5.99	8.95	31	4.42	24.56	17	12.3%	-10.41 [-22.50, 1.68]	
Subtotal (95% CI)			31			17	12.3%	-10.41 [-22.50, 1.68]	
Heterogeneity: Not ap	oplicable								
Test for overall effect:	Z=1.69	(P = 0.)	09)						
5.2.4 Martial Arts v C	control								
Hackney 2009	0.77	8.94	13	4.42	24.56	17	11.2%	-3.65 [-16.30, 9.00]	
Nocera 2013	2.2	10.4	15	3.3	12.6	6	13.9%	-1.10 [-12.47, 10.27]	
Subtotal (95% CI)			28			23	25.1%	-2.24 [-10.70, 6.22]	
Heterogeneity: Chi ² =	: 0.09, df	= 1 (P =	0.77);	I ^z = 0%					
Test for overall effect:	Z = 0.52	? (P = 0.)	60)						
Total (95% CI)			87			63	100.0%	-2.31 [-6.55, 1.92]	-
Heterogeneity: Chi² =	4.70, df	= 4 (P =	0.32):	I ² = 159	6				
Test for overall effect:	Z=1.07	(P=0.)	28)						-20 -10 U 10 20
Test for subgroup dif	ferences	Chi²=	2.05, c	lf = 2 (P	= 0.36),	. I ² = 2.3	3%		Favours intervention Favours No Intervention

E.5.2 Occupational therapy

Patient health related quality of life

Quality asse	ssment					Number of p	oatients	Effect	
Number of studies	Design	Risk of bias	Inconsiste ncy	Indirectnes s	Imprecisio n	от	Control	Mean difference (MD) (95% Cl)	Quality
Generic healt	th related qualit	y of life: EQ5D)						
Sturkenboo m 2014	RCT	Not serious ¹	N/A ²	Not serious ³	Serious ⁴	122	63	0.03; 95%CI -0.03 to 0.08	MODERATE
Parkinson's c	lisease health i	elated quality	of life: PDQ 39						
Sturkenboo m 2014	RCT	Not serious ¹	N/A ²	Not serious ³	Serious ⁵	122	63	-1.7; 95%CI -3.9 to 0.5	MODERATE
¹ Low risk of bi ² N/A: Not appl ³ No serious in ⁴ No serious im	as, as assessed licable as only or directness; popu precision; confic	by NICE RCT q ne study contribu lation was as de dence intervals a	uality checklist uted to this analys scribed in review are tight	sis v protocol					

Activities of daily living

Quality asse	ssment					Number of p	atients	Effect		
Number of studies	Design	Risk of bias	Inconsiste ncy	Indirectnes s	Imprecisio n	от	Control	Mean difference (MD) (95% CI)	Quality	
Canadian par	ticipation 3 mc	onths								
Sturkenboo m 2014	RCT	Not serious ¹	N/A ²	Not serious ³	Not serious	122	63	1.2; 95%C: I 0.8 to 1.6	HIGH	
Canadian pa	Canadian participation 6 months									

¹ Low risk of bias, as assessed by NICE RCT quality checklist

 ² NA: Not applicable as only one study contributed to this analysis
 ³ No serious indirectness; population was as described in review protocol
 ⁴ Serious imprecision: Non-significant results
 ⁵ Cl cross the MID of 1.6 points (Peto et al., 2001)

Quality asse	essment					Number of p	oatients	Effect			
Number of studies	Design	Risk of bias	Inconsiste ncy	Indirectnes s	Imprecisio n	от	Control	Mean difference (MD) (95% Cl)	Quality		
Sturkenboo m 2014	RCT	Not serious ¹	N/A ²	Not serious ³	Not serious	122	63	0.9; 95%CI 0.5 to 1.3	HIGH		
Canadian sat	tisfaction 3 mor	nths									
Sturkenboo m 2014	RCT	Not serious ¹	N/A ²	Not serious ³	Not serious	122	63	1.1; 95%CI 0. to 1.5	HIGH		
Canadian sa	itisfaction 6 mo	nths									
Sturkenboo m 2014	RCT	Not serious ¹	N/A ²	Not serious ³	Not serious	122	63	0.9; 95%CI: 0.5 to 1.3	HIGH		
¹ Low risk of bi ² N/A: Not appl ³ No serious in ⁴ No serious in	Low risk of bias, as assessed by NICE RCT quality checklist N/A: Not applicable as only one study contributed to this analysis No serious indirectness; population was as described in review protocol										

* No serious imprecision; confidence intervais are tight

Recreation and leisure participation

Quality asse	ssment					Number of patients		Effect	
Number of studies	Design	Risk of bias	Inconsiste ncy	Indirectnes s	Imprecisio n	от	Control	Mean difference (MD) (95% CI)	Quality
Utrecht proac	tive coping co	mpetence scale	e						
Sturkenboo m 2014	RCT	Not serious ¹	N/A ²	Not serious ³	Serious ⁴	122	63	0.09: 95%CI -0.02 to 1.21	MODERAT E
Utrecht evalu	ation of rehabi	litation participa	ation satisfactio	on scale					
Sturkenboo m	RCT	Not serious ¹	N/A ²	Not serious ³	Serious ⁴	122	63	3.2; 95%CI -0.6 to 6.8	MODERAT E

 ¹ Low risk of bias, as assessed by NICE RCT quality checklist
 ² NA: Not applicable as only one study contributed to this analysis
 ³ No serious indirectness; population was as described in review protocol

⁴ Non-significant results

Quality asse	ssment					Number of patients		Effect	
Number of studies	Design	Risk of bias	Inconsiste ncy	Indirectnes s	Imprecisio n	от	Control	Mean difference (MD) (95% Cl)	Quality
2014									
¹ Low risk of bia ² N/A: Not appl ³ No serious ind ⁴ No serious ind	as, as assessed icable as only or directness; popu	by NICE RCT q ne study contribu lation was as de	uality checklist uted to this analyse escribed in review	sis v protocol					

Fatigue

Quality asse	ssment					Number of p	oatients	Effect	i l		
Number of studies	Design	Risk of bias	Inconsiste ncy	Indirectnes s	Imprecisio n	от	Control	Mean difference (MD) (95% Cl)	Quality		
Fatigue seve	rity assessmen	nt									
Sturkenboo m 2014	RCT	Not serious ¹	N/A ²	Not serious ³	Serious ⁴	122	63	0.1; 95%CI -0.2 to 0.4	MODERAT E		
¹ Low risk of bi ² NA: Not appli ³ No serious in ⁴ No serious in	¹ Low risk of bias, as assessed by NICE RCT quality checklist ² NA: Not applicable as only one study contributed to this analysis ³ No serious indirectness; population was as described in review protocol ⁴ No serious imprecision: confidence intervals are tight										

Depression

Quality asse	essment					Number of p	oatients	Effect	
Number of studies	Design	Risk of bias	Inconsiste ncy	Indirectnes s	Imprecisio n	от	Control	Mean difference (MD) (95% Cl)	Quality
Becks depres	ssion index								
Sturkenboo m 2014	RCT	Not serious ¹	N/A ²	Not serious ³	Serious ⁴	121	62	-1.4; 95%CI -3.0 to 0.3	MODERAT E
¹ Low risk of b	ias, as assessed	by NICE RCT q	uality checklist						

 ¹ Low risk of bias, as assessed by NICE RCT quality checklist
 ² NA: Not applicable as only one study contributed to this analysis
 ³ No serious indirectness; population was as described in review protocol

⁴ Serious imprecision; non-significant results

Quality asse	essment					Number of patients		Effect	
Number of studies	Design	Risk of bias	Inconsiste ncy	Indirectnes s	Imprecisio n	ОТ	Control	Mean difference (MD) (95% Cl)	Quality
2 NI/A . Not one	laabla oo oob. oo	والمتعادية والمعادية والمعادية	to d to this analy	a i a					

² N/A: Not applicable as only one study contributed to this analysis
 ³ No serious indirectness; population was as described in review protocol
 ⁴ No serious imprecision; confidence intervals are tight

Carer quality of life

Quality asse	ssment					Number of p	oatients	Effect			
Number of studies	Design	Risk of bias	Inconsiste ncy	Indirectnes s	Imprecisio n	от	Control	Mean difference (MD) (95% Cl)	Quality		
Carer quality	of life: EQ5D 3	8 month follow-	up								
Sturkenboo m 2014	RCT	Not serious ¹	N/A ²	Not serious ³	Not serious	112	58	0.06; 95%Cl: 0.02 to 0.11	HIGH		
Carer quality	of life EQ5D: 6	6 month follow u	up								
Sturkenboo m 2014	Sturkenboo RCT Not N/A ² Not Serious ³ 104 59 0.04; 95%CI -0.01 to 0.3 MODERAT 2014 2014 Image: Serious and the cost of the cost										
¹ Low risk of bi ² N/A: Not appl	as, as assessed icable as only o	by NICE RCT q ne study contribu	uality checklist uted to this analy	sis							

³ No serious indirectness; population was as described in review protocol ⁴ No serious imprecision; confidence intervals are tight

E.5.3 Speech and language therapy

Speech impairment: Frenchay dysarthria score

Quality asse	ssment					Number of p	atients	Effect	
Number of studies	Design	Risk of bias	Inconsiste ncy	Indirectnes s	Imprecisio n	Therapy	control	Mean difference (95% Cl)	Quality
Johnson (1990)	RCT	Serious ¹	N/A ²	Serious ³	Not serious	6	6	29 (13.66 to 44.34)	LOW
¹ Serious risk o	of hias: eligibility	criteria randomi	sation method ic	oncealment of al	aceho all inadeo	uately described	4		

nod, concealment of allocation, and placebo all inadequately

² N/A: not applicable, as only one study contributed to analysis

³ Serious indirectness: Eligibility criteria of population of interest ill-defined. It is only implied that all patients had PD.

Vocal loudness

Quality asse	essment					Number of p	oatients	Effect	
Number of studies	Design	Risk of bias	Inconsiste ncy	Indirectnes s	Imprecisio n	Therapy	control	Mean difference (95% Cl)	Quality
Monologue	reading								
2 studies: Johnson (1990) Ramig (2001)	RCT	Serious ¹	Serious ⁴	Serious ³	Not serious	29	21	6.17dB (3.57 to 8.77)	VERY LOW
Monologue	reading - 6 mc	onth follow up							
Ramig (2001)	RCT	Serious ¹	N/A ²	Serious ³	Not serious	14	15	3.5dB (0.9 to 6.1)	LOW
Standard pa	ssage reading	J							
2 studies:	RCT	Serious ¹	Serious ⁵	Serious ³	Not serious	20	21	7.18dB (4.65 to 9.71).	VERY LOW

¹ Serious risk of bias: eligibility criteria, randomisation method, concealment of allocation, and placebo all inadequately described

² NA: not applicable, as only one study contributed to analysis

⁴ Serious inconsistency: $I^2 > 40\%$

⁵ Serious inconsistency: I² >40%

³ Serious indirectness: Eligibility criteria of population of interest ill-defined. It is only implied that all patients had PD.

Quality asse	essment					Number of	patients	Effect			
Number of studies	Design	Risk of bias	Inconsiste ncy	Indirectnes s	Imprecisio n	Therapy	control	Mean difference (95% Cl)	Quality		
Johnson (1990)											
Ramig (2001)											
Standard pa	ssage reading	y - 6 month fo	ollow up								
Ramig (2001)	RCT	Serious ¹	N/A ²	Serious ³	Not serious	14	15	4.5dB (95%CI: 1.9 to 7.1)	LOW		
Loudness of	f prolonged 'a	h' sound									
Ramig (2001)	RCT	Serious ¹	N/A ²	Serious ³	Not serious	14	15	12.1 dB (8.9 to 15.4)	LOW		
Loudness of	f prolonged 'a	h' sound - 6	month follow u	qı							
Ramig (2001)	RCT	Serious ¹	N/A ²	Serious ³	Not serious	14	15	9.4 dB (6.2 to 12.6)	LOW		
¹ Serious risk o ² N/A: not appl	Serious risk of bias: eligibility criteria, randomisation method, concealment of allocation, and placebo all inadequately described N/A: not applicable, as only one study contributed to analysis										

³ Serious indirectness: Eligibility criteria of population of interest ill-defined. It is only implied that all patients had PD.

⁴ Serious imprecision: confidence intervals are wide and cross line of no effect. Cochrane group cite 29 point change as potentially clinically meaningful.

⁵ No serious inconsistency; confidence intervals of estimates overlap

Monotonicity

						Number of patients		Effect	
Number of studies	Design	Risk of bias	Inconsiste ncy	Indirectnes s	Imprecisio n	Therapy	control	Mean difference (95% Cl)	Quality
Maximum pi	tch range								
Johnson (1990)	RCT	Serious ⁴	N/A ²	Serious ³	Very serious ⁵	6	6	66Hz (-4.4 to 136.6)	VERY LOW
Maximum vo	olume range								

¹ Serious risk of bias; Poor randomisation method and poor concealment of allocation. Credibility of placebo condition not clear

² NA: not applicable, as only one study contributed to analysis

³ Serious indirectness: Eligibility criteria of population of interest ill-defined. It is only implied that all patients had PD.

⁴ Serious risk of bias: eligibility criteria, randomisation method, adequate concealment of allocation, and adequate placebo all inadequately described

⁵ Very serious imprecision: Non-significant results and very wide CIs

						Number of patients Effect			
Number of studies	Design	Risk of bias	Inconsiste ncy	Indirectnes s	Imprecisio n	Therapy	control	Mean difference (95% Cl)	Quality
Johnson (1990)	RCT	Serious⁴	N/A ²	Serious ³	Not serious	6	6	23.7dB (9.3 to 38.1)	LOW

¹ Serious risk of bias: eligibility criteria, randomisation method, concealment of allocation, and placebo all inadequately described

² N/A: not applicable, as only one study contributed to analysis

³ Serious indirectness: Eligibility criteria of population of interest ill-defined. It is only implied that all patients had PD.

⁴ Serious imprecision: confidence intervals are wide and cross line of no effect. Cochrane group cite 29 point change as potentially clinically meaningful.

Swallowing safety: penetration aspiration

Quality asse	ssment					Number of patients		Effect	
Number of studies	Design	Risk of bias	Inconsiste ncy	Indirectnes s	Imprecisio n	EMST	sham	Mean difference (95% Cl)	Quality
Troche (2010)	RCT	Not serious ¹	N/A ⁴	Not serious ²	Not serious	30	30	-1.23 (-2.23 to -0.23)	HIGH

¹ Serious risk of bias: eligibility criteria, randomisation method, concealment of allocation, and placebo all inadequately described

² NA: not applicable, as only one study contributed to analysis

³ Serious indirectness: Eligibility criteria of population of interest ill-defined. It is only implied that all patients had PD.

⁴ Serious imprecision: confidence intervals are wide and cross line of no effect. Cochrane group cite 29 point change as potentially clinically meaningful.

Swallowing mechanism: duration of hyoid elevation (s)

Quality asse	ssment					Number of patients		Effect	
Number of studies	Design	Risk of bias	Inconsiste ncy	Indirectnes s	Imprecisio n	EMST	Sham	Mean difference (95% Cl)	Quality
Troche (2010)	RCT	Not serious ¹	N/A ⁴	Not serious ²	serious ³	30	30	0.07s (-4.69 to 4.83)	MODERAT E

¹ Serious risk of bias: eligibility criteria, randomisation method, concealment of allocation, and placebo all inadequately described

² N/A: not applicable, as only one study contributed to analysis

³ Serious indirectness: Eligibility criteria of population of interest ill-defined. It is only implied that all patients had PD.

⁴ Serious imprecision: confidence intervals are wide and cross line of no effect. Cochrane group cite 29 point change as potentially clinically meaningful.

¹ Low risk of bias, as assessed by NICE RCT quality checklist

² No serious indirectness: population clearly defined and match that outlined in review protocol

³ Serious imprecision: non-significant results

Health related quality of life

Quality asse	essment					Number of p	atients	Effect	
Number of studies	Design	Risk of bias	Inconsiste ncy	Indirectnes s	Imprecisio n	EMST	sham	ANOVA F score, p value	Quality
Troche	RCT	Not serious ¹	N/A ¹	Not serious ²	Not serious	30	30	F=3.007 (p=0.007)	LOW
¹ Serious risk o ² N/A: not appl	of bias: eligibility icable. as only o	criteria, randomi ne studv contribu	sation method, c uted to analvsis	concealment of a	llocation, and pl	acebo all inadeo	uately described	1	

¹ NA: not applicable, as only one study contributed to analysis

E.5.4 Nutrition

Low protein diet

Quality asse	essment					No of patients	5		
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Intervention	Control	Effect size (95% CI)	Quality
Total "on" tim	e (Barichel	la 2006)							
1 (Barichella 2006)	RCT	Very serious ^{1,2,3}	N/A	Not serious	Not serious	18	18	MD 114 (19.92 to 208.08)	LOW
Postprandial	"on" time (E	Barichella 2006	3)						
1 (Barichella 2006)	RCT	Very serious ^{1,2,3}	N/A	Not serious	Serious ⁴	18	18	MD 30 (-17.04 to 77.0)	VERY LOW
Total "off" tim	e (Barichel	la 2006)							
1 (Barichella 2006)	RCT	Very serious ^{1,2,3}	N/A	Not serious	Not serious	18	18	MD -107 (-212.53 to -1.47)	LOW
Postprandial	"off" time (E	Barichella 2006	6)						
1 (Barichella 2006)	RCT	Very serious ^{1,2,3}	N/A	Not serious	Serious ⁴	18	18	MD -30 (-77.37 to 17.37)	VERY LOW
1. Uncle 2. Inade 3. Outce 4. Serio	ear if alloca equate blinc omes self-r ous imprecis	tion concealed ling or no blind eported sion: Non-signi	ling ficant result						

Low protein redistribution diet

Quality asse	essment					No of patients				
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Intervention	Control	Effect size (95% CI)	Quality	
Percentage of	of "on" time									
1 (Tsul 1989)	RCT	Very serious ^{1,2,3,4}	N/A	Not serious	Serious ⁵	10	10	MD 10.55 (-4.28, 25.58)	VERY LOW	
Modified Col	umbia score	es								
1 (Tsul 1989)	RCT	Very serious ^{1,2,3,4}	N/A	Not serious	Serious ⁵	10	10	MD -3.98 (-14.82, 6.86)	VERY LOW	
Total "off" tim	ne									
1 (Croxson 1991)	RCT	Very serious ^{2,6}	N/A	Not serious	Serious ⁵	8	8	MD -0.81 (-6.23, 4.61)	VERY LOW	
1. Uncle	ear method	of randomisat	ion							
2. Uncl	ear if alloca	tion concealed	l							
3. No p	3. No precise definition of outcome									
4. Inap	propriate ler	ngth of follow-u	qu							
5. Serio	5. Serious imprecision: Non-significant result									
6. Outc	omes self-r	eported								

Comparison of two low-protein diets

Quality asse	uality assessment						5		
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Intervention	Control	Effect size (95% CI)	Quality
Time spent in									
1 (Barichella 2007)	RCT	Very serious ^{1,2,3,4}	N/A	Not serious	Serious ⁵	6	6	MD 0.37 (-1.13, 1.87)	VERY LOW
Energy exper	nditure								
1 (Barichella 2007)	RCT	Very serious ^{1,2,3,4}	N/A	Not serious	Serious ⁵	6	6	MD 172 (-128, 472)	VERY LOW

Quality asse	ssment					No of patients				
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Intervention	Control	Effect size (95% CI)	Quality	
Patient globa	l improveme	ent (very much	better/much bette	er)						
1 (Barichella 2007)	RCT	Very serious ^{1,2,3,4}	N/A	Not serious	Serious ⁵	6	6	RR 13.00 (0.89, 189.39)	VERY LOW	
Patient globa	l improveme	ent (very much	better/much bette	er)						
1 (Barichella 2007)	RCT	Very serious ^{1,2,3,4}	N/A	Not serious	Serious ⁵	6	6	RR 0.08 (0.01, 1.12)	VERY LOW	
 Uncle Inade Inapp Outco Serio 	 Unclear if allocation concealed Inadequate blinding or no blinding Inappropriate length of follow up Outcomes self reported Serious imprecision: Non-significant result 									

High fibre supplement

Quality asse	essment					No of patients			
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Intervention	Control	Effect size (95% CI)	Quality
Absorption: a	irea under t	he curve							
1 (Fernandez -Martinez 2014)	RCT	Serious ¹	N/A	Not serious	Serious ²	18	18	MD -0.63 (-10.30, 9.04)	LOW
Absorption: p	eak plasma	a concentration	ו						
1 (Fernandez -Martinez 2014)	RCT	Serious ¹	N/A	Not serious	Serious ²	18	18	MD -64.20 (-184.92, 56.52)	LOW
Absorption: ti	me to peak	blood level							

Quality asse	ssment					No of patients	6		
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Intervention	Control	Effect size (95% CI)	Quality
1 (Fernandez -Martinez 2014)	RCT	Serious ¹	N/A	Not serious	Serious ²	18	18	MD 3.55 (-10.96, 18.06)	LOW
1. Uncle 2. Serio	ear if alloca	tion concealed sion: non-signif	ficant results						

Fasting diet

Quality asse	essment					No of patients			
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Intervention	Control	Effect size (95% CI)	Quality
Absorption: a	rea under t	he curve							
1 (Brefel 1998)	RCT	Very serious ^{1,2,3}	N/A	Not serious	Serious ⁴	12	12	MD 3.20 (-4.93, 11.33)	VERY LOW
Absorption: p	eak plasma	a concentration	ו						
1 (Brefel 1998)	RCT	Very serious ^{1,2,3}	N/A	Not serious	Serious ⁴	12	12	MD 1.52 (-0.16, 3.20)	VERY LOW
Absorption: ti	me to peak	plasma conce	entration						
1 (Brefel 1998)	RCT	Very serious ^{1,2,3}	N/A	Not serious	Serious ⁵	12	12	MD -2.12 (-2.81, -1.43)	VERY LOW
 Uncle Uncle Uncle Inade Serio 	ear method ear if alloca equate blinc ous imprecis	of randomisat tion concealed ling or no blinc sion: non-signit	ion l ling ficant results						

5. Means and SD imputed from medians and ranges

Creatine supplementation

Quality asse	essment					No of patients			
No of	Decian	Risk of	Inconsistency	Indirectocoo	Improcision	Intervention	Control	Effect size (0.5% CI)	Quality
Studies	Design	Dias	n difference from		imprecision		Control	Effect Size (95% CI)	Quality
SF-30 Gener		A contraction (mea		baseline) - belle			00		
1 (Bender 2006)	RCI	very serious ²	N/A	Not serious	Serious	40	20	MD 5.00 (-4.53, 14.53)	LOW
SF-36 Vitality	/ (mean diff	erence from ba	aseline) - better in	dicated by lower	values				
1 (Bender 2006)	RCT	Very serious ²	N/A	Not serious	Serious ³	40	20	MD 7.00 (-1.43, 15.43)	VERY LOW
SF-36 Role li	mitations (r	mean differenc	e from baseline) -	better indicated	by lower values				
1 (Bender 2006)	RCT	Very serious ²	N/A	Not serious	Not serious	40	20	MD 21.00 (5.29, 36.70)	LOW
SF-36 Menta	l health (me	ean difference	from baseline) - b	etter indicated b	y lower values				
1 (Bender 2006)	RCT	Very serious ²	N/A	Not serious	Not serious	40	20	MD 8.00 (0.03, 15.97)	LOW
SF-36 Social	functioning	(mean differe	nce from baseline) - better indicat	ed by lower valu	es			
1 (Bender 2006)	RCT	Very serious ²	N/A	Not serious	Serious ³	40	20	MD 4.00 (-5.62, 13.62)	VERY LOW
SF-36 Bodily	pain (mear	n difference fro	m baseline) - bett	er indicated by l	ower values				
1 (Bender 2006)	RCT	Very serious ²	N/A	Not serious	Serious ³	40	20	MD -6.00 (-21.12, 9.12)	VERY LOW
SF-36 Role li	mitations (r	mean differenc	e from baseline) -	better indicated	by lower values	i			
1 (Bender 2006)	RCT	Very serious ²	N/A	Not serious	Serious ³	40	20	MD -10.00 (-30.32, 10.32)	VERY LOW
SF-36 Physic	cal functioni	ng (mean diffe	rence from baseli	ne) - better indic	ated by lower va	alues			
1 (Bender 2006)	RCT	Very serious ²	N/A	Not serious	Serious ³	40	20	MD -4.00 (-14.08, 6.08)	VERY LOW
EQ-5D (mea	n difference	from baseline)						
1 (Kieburtz 2015)	RCT	Not serious	N/A	Not serious	Serious ³	334	342	MD 0.005 (-0.03, 0.04)	MODE RATE

Quality asse	essment					No of patients	;		
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Intervention	Control	Effect size (95% CI)	Quality
PDQ-39 Sum	mary Index	(mean differe	nce from baseline)					
1 (Kieburtz 2015)	RCT	Not serious	N/A	Not serious	Very serious ⁶	477	478	MD -1.2 (-4.2, 1.7)	LOW
Total UPDRS	score (me	an difference f	rom baseline) - be	etter indicated by	lower values				
2 (Bender 2006, Kieburtz 2015)	RCT	Serious ¹	Not serious	Not serious	Serious ³	370	356	MD 1.02 (-1.11, 3.15)	LOW
Total UPDRS	score (me	an difference f	rom baseline) - be	etter indicated by	lower values				
1 (Haas 2007)	RCT	Serious ¹	N/A	Not serious	Serious ³	10	10	MD -1.70 (-7.08, 3.68)	LOW
UPDRS com	plications (r	mean differenc	e from baseline) -	better indicated	by lower values				
1 (Bender 2006)	RCT	Very serious ²	N/A	Not serious	Serious ³	40	20	MD 0.20 (-0.55, 0.95)	VERY LOW
UPDRS moto	or score (me	ean difference	from baseline) - b	etter indicated b	y lower values				
2 (Bender 2006, Kieburtz 2015)	RCT	Serious ¹	Not serious	Not serious	Not serious	370	356	MD 0.44 (-1.02, 1.90)	MODE RATE
UPDRS moto	or score (me	ean difference	from baseline) - b	etter indicated b	y lower values				
1 (Haas 2007)	RCT	Serious ¹	N/A	Not serious	Serious ⁶	10	10	MD -3.90 (-8.03, 0.23)	LOW
UPDRS ADL	score (mea	an difference fr	om baseline) - bet	tter indicated by	lower values				
2 (Bender 2006, Kieburtz 2015)	RCT	Serious ¹	Not serious	Not serious	Serious ⁴	373	359	MD 1.30 (-1.12, 3.72)	LOW
UPDRS ADL	score (mea	an difference fr	om baseline) - bei	tter indicated by	lower values				
1 (Haas 2007)	RCT	Serious ¹	N/A	Not serious	Not serious	10	10	MD -0.20 (-2.20, 1.80)	MODE RATE

Quality asse	ssment					No of patients	;		
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Intervention	Control	Effect size (95% CI)	Quality
UPDRS ment	tal health so	core (mean diff	ference from base	line) - better ind	icated by lower	values			
2 (Bender 2006, Kieburtz 2015)	RCT	Serious ¹	Serious⁵	Not serious	Serious ³	373	359	MD -0.0 (-0.27, 0.26)	VERY LOW
UPDRS ment	tal health so	core (mean dif	ference from base	line) - better ind	icated by lower	values			
1 (Haas 2007)	RCT	Serious ¹	N/A	Not serious	Serious ³	10	10	MD 0.40 (-0.08, 0.88)	LOW
Hoehn and Y	ahr score (mean differend	e from baseline) -	better indicated	by lower values	6			
1 (Haas 2007)	RCT	Serious ¹	N/A	Not serious	Not serious	10	10	MD -0.40 (-0.58, -0.22)	MODE RATE
Mass, kg (me	an differen	ce from baseli	ne) - better indicat	ed by lower valu	les				
1 (Haas 2007)	RCT	Serious ¹	N/A	Not serious	Serious ³	10	10	MD 0.40 (-4.74, 5.54)	LOW
BMI (mean di	ifference fro	om baseline)							
1 (Kieburtz 2015)	RCT	Not serious	N/A	Not serious	Serious ³	338	341	MD -0.3 (-0.8, 0.2)	MODE RATE
Levodopa do	se (mean d	ifference from	baseline) - better	indicated by low	er values				
1 (Bender 2006)	RCT	Very serious ²	N/A	Not serious	Serious ³	40	20	MD -57.0 (-145.3, 31.3)	VERY LOW
Dopamine ag	jonist dose	(mean differer	nce from baseline)	- better indicate	d by lower value	s			
1 (Bender 2006)	RCT	Very serious ²	N/A	Not serious	Not serious	40	20	MD -132.0 (-195.8, -68.3)	LOW
 Seric Very Non- CI cr 1² > 4 CI cr 	ous risk of b serious risl significant r oss the MIE 0% oss the MIE	ias detected of bias detect esult of 3 (Schrag	ted et al., 2006) 2001)						

Amino acid supplementation

Quality asse	essment					No of patients			
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Intervention	Control	Effect size (95% CI)	Quality
Body weight	(kg) (mean	difference from	n baseline) - bette	r indicated by lo	wer values				
1 (Cucca 2015)	RCT	Serious ¹	N/A	Not serious	Serious ²	7	7	MD -6.50 (-13.71, 0.71)	LOW
Body motor s	score (kg) (r	mean differenc	e from baseline) -	better indicated	by lower values				
1 (Cucca 2015)	RCT	Serious ¹	N/A	Not serious	Very serious ³	7	7	MD 3.20 (-3.60, 10.00)	VERY LOW
 Seric Non- CI cr 	ous risk of b significant r oss the MIE	ias detected result D of 3.25 (Horv	ath et al., 2015)						

Co-enzyme Q10 supplemention

Quality asse	essment					No of patients	5		
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Intervention	Control	Effect size (95% CI)	Quality
Total UPDRS	6 (mean diff	erence from ba	aseline) - better in	dicated by lower	r values				
3 (Negida 2016)	SR	Not serious	Not serious	Not serious	Serious ¹	475	468	MD -0.05 (-0.25, 0.15)	MODE RATE
UPDRS moto	or score (me	ean difference	from baseline) - b	etter indicated b	y lower values				
4 (Negida 2016)	SR	Not serious	Not serious	Not serious	Not serious	546	539	MD 0.05 (-0.07, 0.17)	HIGH
UPDRS ADL	score (mea	an difference fr	om baseline) - be	tter indicated by	lower values				
4 (Negida 2016)	SR	Not serious	Serious ²	Not serious	Not serious	546	539	MD -0.10 (-0.35, 0.15)	MODE RATE
UPDRS men	tal health so	core (mean dif	ference from base	line) - better ind	icated by lower	values			
4 (Negida 2016)	SR	Not serious	Not serious	Not serious	Serious ¹	546	539	MD -0.03 (-0.23, 0.17)	MODE RATE
Schwab and	England mo	odified score (r	mean difference fr	om baseline) - b	better indicated b	v lower values			

Quality asse	essment					No of patients	5		
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Intervention	Control	Effect size (95% CI)	Quality
3 (Negida 2016)	SR	Not serious	Serious ²	Not serious	Serious ¹	546	539	MD 0.08 (-0.17, 0.29)	LOW
UPDRS ADL	/motor scor	e (mean differ	ence from baseline	e) - better indica	ted by lower val	ues			
1 (Storch 2007)	RCT	Not serious	N/A	Not serious	Serious ¹	64	67	MD 2.15 (-1.08, 5.38)	MODE RATE
1. Non- 2. Cons	significant r siderable be	result etween study h	eterogeneity						

Trigonella foenum-gracum I seeds

Quality asse	essment					No of patients	;				
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Intervention	Control	Effect size (95% CI)	Quality		
Total UPDRS	(mean diff	erence from ba	aseline) - better in	dicated by lower	r values						
1 (Nathan 2014)	RCT	Not serious	N/A	Not serious	Serious ¹	23	19	MD -5.36 (-13.70, 2.98)	MODE RATE		
UPDRS moto	PDRS motor score (mean difference from baseline) - better indicated by lower values										
1 (Nathan 2014)	RCT	Not serious	N/A	Not serious	Serious ²	23	19	MD -4.76 (-11.82, 2.30)	MODE RATE		
UPDRS ADL	score (mea	an difference fr	om baseline) - bei	tter indicated by	lower values						
1 (Nathan 2014)	RCT	Not serious	N/A	Not serious	Very serious ²	23	19	MD 0.07 (-3.66, 3.80)	LOW		
UPDRS ment	tal health so	core (mean dif	ference from base	line) - better ind	icated by lower	values					
1 (Nathan 2014)	RCT	Not serious	N/A	Not serious	Serious ¹	23	19	MD -0.65 (-2.03, 0.73)	MODE RATE		
Hoehn and Y	ahr Stage r	eversal									
1 (Nathan 2014)	RCT	Not serious	N/A	Not serious	Serious ¹	23	19	RR 4.13 (0.53, 32.38)	MODE RATE		
Hoehn and Y	ahr Stage u	unchanged									

Quality asse	essment					No of patients	5		
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Intervention	Control	Effect size (95% CI)	Quality
1 (Nathan 2014)	RCT	Not serious	N/A	Not serious	Serious ¹	23	19	RR 0.83 (0.57, 1.21)	MODE RATE
Hoehn and Y	ahr Stage a	advancement							
1 (Nathan 2014)	RCT	Not serious	N/A	Not serious	Serious ¹	23	19	RR 0.83 (0.19, 3.63)	MODE RATE
1. Non- 2. Cl cr	significant r	esult) of 3.25 (Horv	ath et al., 2015)						

Vitamin D supplementation

Quality asse	uality assessment						;		
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Intervention	Control	Effect size (95% CI)	Quality
PDQ39 total	(mean diffe	rence from bas	seline) - better ind	icated by lower	values				
1 (Suzuki 2013)	RCT	Not serious	N/A	Not serious	Serious ¹	55	57	MD 2.26 (-8.72, 4.20)	MODE RATE
PDQ39 cogn	itive impairr	ment (mean dif	ference from base	eline) - better inc	licated by lower	values			
1 (Suzuki 2013)	RCT	Not serious	N/A	Not serious	Serious ²	55	57	MD -1.50 (-8.08, 5.08)	MODE RATE
PDQ39 socia	al support (r	nean difference	e from baseline) -	better indicated	by lower values				
1 (Suzuki 2013)	RCT	Not serious	N/A	Not serious	Serious ²	55	57	MD -3.65 (-10.53, 3.23)	MODE RATE
PDQ39 bodil	y support (r	nean differenc	e from baseline) -	better indicated	by lower values				
1 (Suzuki 2013)	RCT	Not serious	N/A	Not serious	Serious ²	55	57	MD -5.67 (-13.63, 2.29)	MODE RATE
PDQ39 comr	nunication	(mean differen	ce from baseline)	- better indicate	d by lower value	S			
1 (Suzuki 2013)	RCT	Not serious	N/A	Not serious	Serious ²	55	57	MD -2.17 (-9.70, 5.36)	MODE RATE
PDQ39 stigm	na (mean di	fference from b	baseline) - better i	ndicated by lowe	er values				

Quality asse	essment					No of patients			
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Intervention	Control	Effect size (95% CI)	Quality
1 (Suzuki 2013)	RCT	Not serious	N/A	Not serious	Serious ²	55	57	MD 5.75 (-1.88, 13.38)	MODE RATE
PDQ39 emot	ional wellbe	eing (mean diff	erence from base	line) - better indi	cated by lower	alues			
1 (Suzuki 2013)	RCT	Not serious	N/A	Not serious	Serious ²	55	57	MD -1.71 (-9.94, 6.52)	MODE RATE
PDQ39 activi	ities of daily	living (mean c	lifference from bas	seline) - better ir	ndicated by lowe	er values			
1 (Suzuki 2013)	RCT	Not serious	N/A	Not serious	Serious ²	55	57	MD -1.64 (-10.64, 7.36)	MODE RATE
PDQ39 mobi	lity (mean d	lifference from	baseline) - better	indicated by low	ver values				
1 (Suzuki 2013)	RCT	Not serious	N/A	Not serious	Serious ²	55	57	MD -3.03 (-12.62, 6.56)	MODE RATE
EQ-5D (mean	n difference	from baseline) - better indicated	d by lower value	S				
1 (Suzuki 2013)	RCT	Not serious	N/A	Not serious	Serious ²	55	57	MD 0.05 (-0.05, 0.15)	MODE RATE
MMSE (Stage	e 1-5) - bett	ter indicated by	/ lower values						
1 (Suzuki 2013)	RCT	Not serious	N/A	Not serious	Serious ²	55	57	MD -0.60 (-1.33, 0.13)	MODE RATE
Total UPDRS	6 (mean diff	erence from ba	aseline) - better in	dicated by lower	r values				
1 (Suzuki 2013)	RCT	Not serious	N/A	Not serious	Not serious	55	57	MD -5.07 (-10.13, -0.01)	HIGH
UPDRS com	plications (r	nean differenc	e from baseline) -	better indicated	by lower values				
1 (Suzuki 2013)	RCT	Not serious	N/A	Not serious	Serious ²	55	57	MD -0.09 (-0.62, 0.44)	MODE RATE
UPDRS moto	or score (me	ean difference	from baseline) - b	etter indicated b	y lower values				
1 (Suzuki 2013)	RCT	Not serious	N/A	Not serious	Serious ³	55	57	MD -2.10 (-5.64, 1.44)	MODE RATE
UPDRS ADL	score (mea	an difference fr	om baseline) - be	tter indicated by	lower values				
1 (Suzuki 2013)	RCT	Not serious	N/A	Not serious	Serious ³	55	57	MD -5.24 (-10.32, -0.16)	MODE RATE

Quality asse	uality assessment						5		
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Intervention	Control	Effect size (95% CI)	Quality
UPDRS men	tal health so	core (mean dif	ference from base	line) - better ind	icated by lower	values			
1 (Suzuki 2013)	RCT	Not serious	N/A	Not serious	Serious ²	55	57	MD -0.38 (-0.93, 0.17)	MODE RATE
Hoehn & Yah	nr score (me	ean difference	from baseline) - b	etter indicated b	y lower values				
1 (Suzuki 2013)	RCT	Not serious	N/A	Not serious	Not serious	55	57	MD -0.31 (-0.55, 0.07)	HIGH
1. Clcr 2. Non- 3. Clcr 4. Clcr	oss the MIE significant r oss the MIE oss the MIE) of 1.6 (Peto e result) of 3.25 (Horv) of 3 (Schrag	et al., 2001) ath et al., 2015) et al., 2006)						

Forest plots (CoQ10 supplementation)

UPDRS Total

		CoQ10			Placebo		1	Std. Mean Difference	Std. Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Random, 95% C	IV, Random, 95% CI
1.1.1 dose < 500 mg/c	day								
Muller 2003	-2.29	2.92	14	-1.21	2.64	14	6.2%	-0.38 [-1.13, 0.37]	
QE2 2002	8.81	7.925986	21	11.9	8.163265	16	7.8%	-0.38 [-1.03, 0.28]	
Subtotal (95% CI)			35			30	14.0%	-0.38 [-0.87, 0.12]	-
Heterogeneity: Tau ² =	0.00; Cł	hi ² = 0.00, d	f = 1 (P	= 1.00); I ² = 0%				
Test for overall effect:	Z = 1.50	(P = 0.13)							
1.1.2 dose = 600 mg/d	day								
QE2 2002	10.82	7.837646	20	11.9	8.163265	16	7.8%	-0.13 [-0.79, 0.53]	
Subtotal (95% CI)			20			16	7.8%	-0.13 [-0.79, 0.53]	
Heterogeneity: Not app	plicable								
Test for overall effect:	Z = 0.39	P = 0.69							
1.1.3 dose = 1200 mg	/day								
QE2 2002	6.69	7.829929	23	11.9	8.163265	16	7.8%	-0.64 [-1.30, 0.01]	
QE3 2014	7.5	8.790017	201	6.92	8.976118	203	35.3%	0.07 [-0.13, 0.26]	
Subtotal (95% CI)			224			219	43.1%	-0.22 [-0.89, 0.46]	
Heterogeneity: Tau ² =	0.19; Cl	hi ² = 4.09, d	f = 1 (P	= 0.04); l ² = 76%				
Test for overall effect:	Z = 0.62	? (P = 0.53)							
1.1.4 dose = 2400 mg	/day								
QE3 2014	8.01	8.82	196	6.92	8.976118	203	35.1%	0.12 [-0.07, 0.32]	
Subtotal (95% CI)			196			203	35.1%	0.12 [-0.07, 0.32]	•
Heterogeneity: Not app	plicable								
Test for overall effect:	Z = 1.22	? (P = 0.22)							
N									
Total (95% CI)			475			468	100.0%	-0.05 [-0.25, 0.15]	
Heterogeneity: Tau ² =	0.02; Cl	hi ^z = 7.83, d	f = 5 (P	= 0.17); l ^a = 36%				2 1 0 1 2
Test for overall effect:	Z = 0.47	(P = 0.64)							Favours CoQ10 Favours Placebo
Test for subgroup diffe	rences:	Chi ² = 4.19	df = 3	(P = 0.3)	24), P = 28.	4%			

UPDRSI(mental)

		CoQ10			Placebo		1	Std. Mean Difference	Std. Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Random, 95% CI	IV, Random, 95% CI
1.2.1 dose < 500 mg/	day								
QE2 2002	0.54	0.94691	21	0.9	0.969388	16	7.6%	-0.37 [-1.02, 0.29]	
Subtotal (95% CI)			21			16	7.6%	-0.37 [-1.02, 0.29]	
Heterogeneity: Not ap	plicable								
Test for overall effect:	Z = 1.10	(P = 0.27)							
1.2.2 dose = 600 mg/	day								
QE2 2002	0.35	0.946906	20	0.9	0.969388	16	7.3%	-0.56 [-1.23, 0.11]	
Subtotal (95% CI)			20			16	7.3%	-0.56 [-1.23, 0.11]	
Heterogeneity: Not ap	plicable								
Test for overall effect:	Z = 1.64	(P = 0.10)							
1.2.3 dose = 1200 mg	g/day								
QE2 2002	0.33	0.942038	23	0.9	0.969388	16	7.7%	-0.59 [-1.24, 0.07]	
QE3 2014	0.45	1.27597	201	0.41	1.282303	203	29.1%	0.03 [-0.16, 0.23]	
Subtotal (95% CI)			224			219	36.8%	-0.20 [-0.78, 0.39]	
Heterogeneity: Tau ² =	0.13; CI	hi ² = 3.15, d	f = 1 (F	P = 0.08); l ² = 68%				
Test for overall effect:	Z = 0.66	6 (P = 0.51)							
1.2.4 dose = 2400 mg	g/day								
Net-PD 2007	0.68	1.49	71	0.48	1.18	71	19.3%	0.15 [-0.18, 0.48]	
QE3 2014	0.6	1.26	196	0.41	1.282303	203	29.0%	0.15 [-0.05, 0.35]	-
Subtotal (95% CI)			267			274	48.3%	0.15 [-0.02, 0.32]	•
Heterogeneity: Tau ² =	0.00; CI	hi ² = 0.00, d	f = 1 (F	P = 1.00); I ² = 0%				
Test for overall effect:	Z = 1.73	3 (P = 0.08)							
Total (95% CI)			532			525	100.0%	-0.03 [-0.23, 0.17]	+
Heterogeneity: Tau ² =	0.03; CI	hi ² = 9.71, d	f = 5 (F	= 0.08); l ² = 48%			_	
Test for overall effect:	Z = 0.32	P = 0.75							-1 -0.5 0 0.5 1 Eavours CoO10 Eavours Placebo
Test for subgroup diffe	erences:	Chi2 = 6.73	, df = 3	(P = 0)	08), l ² = 55.	4%			Favours Courto Favours Placebo

UPDRS II (ADL)

		CoQ10			Placebo		:	Std. Mean Difference	Std. Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Random, 95% CI	IV, Random, 95% CI
1.3.1 dose < 500 mg/s	day								
QE2 2002	2.54	3.273268	21	4.74	3.346939	16	9.5%	-0.65 [-1.32, 0.02]	
Subtotal (95% CI)			21			16	9.5%	-0.65 [-1.32, 0.02]	
Heterogeneity: Not app	plicable								
Test for overall effect:	Z = 1.91	(P = 0.06)							
1.3.2 dose = 600 mg/	day								
QE2 2002	4.02	3.240017	20	4,74	3.346939	16	9.7%	-0.21 [-0.87, 0.45]	
Subtotal (95% CI)			20			16	9.7%	-0.21 [-0.87, 0.45]	
Heterogeneity: Not app	plicable								
Test for overall effect:	Z = 0.64	(P = 0.52)							
1.3.3 dose = 1200 mg	/day								
QE2 2002	1.62	3.217611	23	4.74	3.346939	16	9.4%	-0.93 [-1.61, -0.26]	
QE3 2014	2.76	3.260813	201	2.23	3.276996	203	25.8%	0.16 [-0.03, 0.36]	-
Subtotal (95% CI)			224			219	35.2%	-0.34 [-1.41, 0.73]	
Heterogeneity: Tau ² =	0.54; Cł	hi ² = 9.36, d	f = 1 (F	P = 0.00	2); l ² = 89%				
Test for overall effect:	Z = 0.62	2 (P = 0.54)							
1.3.4 dose = 2400 mg	/day								
Net-PD 2007	2.15	3.05	71	2.04	2.93	71	19.9%	0.04 [-0.29, 0.37]	
QE3 2014	2.5	3.22	196	2.23	3.276996	203	25.7%	0.08 [-0.11, 0.28]	
Subtotal (95% CI)			267			274	45.7%	0.07 [-0.10, 0.24]	*
Heterogeneity: Tau ² =	0.00; Cł	hi ² = 0.06, d	f = 1 (F	P = 0.81); I ² = 0%				
Test for overall effect:	Z = 0.82	2 (P = 0.41)							
Total (95% CI)			532			525	100.0%	-0.10 [-0.35, 0.15]	-
Heterogeneity: Tau ² =	0.05; Cł	hi ² = 14.37,	df = 5	(P = 0.0	1); l ² = 65%			+	
Test for overall effect:	Z = 0.79	(P=0.43)						-2	-1 U 1
Test for subgroup diffe	rences:	Chi ² = 5.10	. df = 3	(P = 0.	16), l ² = 41,	1%			Favours courte Favours Placebo

UPDRS III (motor)

		CoQ10			Placebo			Std. Mean Difference	Std. Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Random, 95% CI	IV, Random, 95% CI
.4.1 dose < 500 mg/	day								
Muller 2003	-0.29	0.61	14	-0.64	1.28	14	2.5%	0.34 [-0.41, 1.09]	
QE2 2002	5.88	5.856812	21	6.54	6.071429	16	3.4%	-0.11 [-0.76, 0.54]	
Subtotal (95% CI)			35			30	5.9%	0.08 [-0.41, 0.58]	
Heterogeneity: Tau ² =	0.00; Cł	hi ² = 0.78, d	f = 1 (F	= 0.38); I ² = 0%				
Test for overall effect:	Z = 0.34	(P = 0.74)							
1.4.2 dose = 600 mg/	day								
QE2 2002	6.47	5.795523	20	6.54	6.071429	16	3.3%	-0.01 [-0.67, 0.65]	
Subtotal (95% CI)			20			16	3.3%	-0.01 [-0.67, 0.65]	
Heterogeneity: Not ap	plicable								
Test for overall effect:	Z = 0.03	3 (P = 0.97)							
1.4.3 dose = 1200 mg	/day								
QE2 2002	4.61	5.786807	23	6.54	6.071429	16	3.4%	-0.32 [-0.96, 0.32]	
QE3 2014	4.24	6.379851	201	4.23	6.411513	203	37.4%	0.00 [-0.19, 0.20]	
Subtotal (95% CI)			224			219	40.8%	-0.03 [-0.21, 0.16]	•
Heterogeneity: Tau ² =	0.00; Cł	hi ² = 0.88, d	f = 1 (F	= 0.35); I ² = 0%				
Test for overall effect:	Z = 0.27	(P = 0.79)							
1.4.4 dose = 2400 mg	/day								
Net-PD 2007	4.73	6.66	71	3.79	6.16	71	13.1%	0.15 [-0.18, 0.48]	
QE3 2014	4.88	6.3	196	4.23	6.411513	203	36.9%	0.10 [-0.09, 0.30]	+
Subtotal (95% CI)			267			274	50.0%	0.11 [-0.06, 0.28]	*
Heterogeneity: Tau ^a =	0.00; Cł	hi ² = 0.05, d	f = 1 (F	= 0.82); I ² = 0%				
Test for overall effect:	Z = 1.32	t (P = 0.19)							
Total (95% CI)			546			539	100.0%	0.05 [-0.07, 0.17]	•
Heterogeneity: Tau ² =	0.00; Cł	hi ² = 2.94, d	f = 6 (P	= 0.82); I ² = 0%			-	
Test for overall effect:	Z = 0.84	(P = 0.40)							-1 -0.5 0 0.5 1
Test for subgroup diffe	rences:	Chi ² = 1.23	. df = 3	(P = 0.1)	75), I ² = 0%				Payours Courto Payours Placebo

Schwab and England modified score ("for examiner")

		CoQ10			Placebo			Std. Mean Difference	Std. Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Random, 95% CI	IV, Random, 95% CI
1.5.1 dose < 500 mg/c	iay								
QE2 2002	-4.89	5.120327	21	-7.98	5.234694	16	7.9%	0.58 [-0.08, 1.25]	
Subtotal (95% CI)			21			16	7.9%	0.58 [-0.08, 1.25]	
Heterogeneity: Not app	licable								
Test for overall effect:	Z = 1.72	(P = 0.09)							
1.5.2 dose = 600 mg/c	fay								
QE2 2002	-7.03	5.099604	20	-7.98	5.234694	16	8.1%	0.18 [-0.48, 0.84]	
Subtotal (95% CI)			20			16	8.1%	0.18 [-0.48, 0.84]	
Heterogeneity: Not app	licable								
Test for overall effect: 2	Z = 0.54	(P = 0.59)							
1.5.3 dose = 1200 mg	/day								
QE2 2002	-3.55	5.07722	23	-7.98	5.234694	16	7.9%	0.84 [0.18, 1.51]	
QE3 2014	-4.29	8.790017	201	-4.07	8.83364	203	28.3%	-0.02 [-0.22, 0.17]	-
Subtotal (95% CI)			224			219	36.2%	0.35 [-0.49, 1.19]	
Heterogeneity: Tau ² =	0.31; Cł	h ² = 5.99, d	f = 1 (P	e = 0.01); l² = 83%				
Test for overall effect:	Z = 0.81	(P = 0.42)							
1.5.4 dose = 2400 mg	/day								
Net-PD 2007	-5.5	8.16	71	-4.9	7.63	71	19.6%	-0.08 [-0.40, 0.25]	
QE3 2014	-4.94	8.68	196	-4.07	8.83364	203	28.2%	-0.10 [-0.30, 0.10]	-
Subtotal (95% CI)			267			274	47.8%	-0.09 [-0.26, 0.08]	•
Heterogeneity: Tau ² =	0.00; Cł	h ² = 0.01, d	f = 1 (P	e = 0.90); l ² = 0%				
Test for overall effect:	Z = 1.08	(P = 0.28)							
Total (95% CI)			532			525	100.0%	0.08 [-0.13, 0.29]	🕈
Heterogeneity: Tau ² =	0.03; Cł	hi ² = 10.63,	df = 5 (P = 0.0	6); I ² = 53%				2 1 0 1 2
Test for overall effect:	Z = 0.72	(P = 0.47)							Favours CoQ10 Favours Placebo
Test for subgroup diffe	rences:	Chi2 = 4.97	. df = 3	(P = 0.1	17), I ² = 39.	7%			

Forest plots (Creatine supplementation)

UPDRS Total



UPDRS Motor



UPDRS ADL



UPDRS Mental



E.6 Advanced therapies: deep brain stimulation and levodopa-carbidopa intestinal gel

- E.6.1 Deep brain stimulation compared with best medical treatment for advanced Parkinson's disease
- E.6.1.1 Adverse events

						No. of events / no. of patients	or patient-		
		Quality asse	essment			years			
No. of		Risk of							
studies	Design	bias	Inconsistency	Indirectness	Imprecision	DBS	Control	Effect (95% CI)	Quality
Serious	adverse e	events (proba	bility of experien	icing ≥1)					
4 ^{1,2,3,4}	RCT	No serious	Not serious ⁵	Serious ⁶	No serious	138/496	48/361	RR = 2.26 (1.57 to 3.23)	MODERATE
Serious	adverse e	events (rate p	er patient-year)						
4 ^{1,2,3,4}	RCT	No serious	Not serious ⁵	Serious ⁶	No serious	208 per 314.25pt-yrs	58 per 291.25pt-yrs	IRR = 2.72 (1.60 to 4.64)	MODERATE
Falls (pr	obability	of experiencir	ng ≥1)						
4 ^{1,2,3,4}	RCT	No serious	Serious ⁷	Serious ⁶	Serious ⁸	29/496	14/361	RR = 1.24 (0.32 to 4.83)	VERY LOW
Falls (rat	te per pat	ient-year)							
4 ^{1,2,3,4}	RCT	No serious	Serious ⁷	Serious ⁶	Serious ⁸	30 per 314.25pt-yrs	14 per 291.25pt-yrs	IRR = 1.44 (0.45 to 4.62)	VERY LOW

¹ Okun 2012

² Deuschl 2006

³ Weaver 2009

⁴ Williams 2010 (main PDSURG publication [all participants regardless of HY score]; no subgroup data available for this outcome)

⁵ Statistical heterogeneity observed; however, this is almost wholly ascribable to differences between Okun 2012 and other studies, and this is explicable on the grounds that participants in the control arm of Okun 2012 underwent surgical implantation of inert device, so not downgraded

⁶ Some RCTs include a nontrivial proportion of participants with less advanced PD than the specified population for this question

⁷ Marked statistical heterogeneity and inconsistency in definition of events (some RCTs report all recorded falls; some falls leading to fracture only)

⁸ At a 95% confidence level, data are consistent with appreciable harm, appreciable benefit and no difference

	DBS	5	Medica	tion		Risk Ratio	Risk Ratio
Study or Subgroup	Events	Total	Events	Total	Weight	M-H, Random, 95% Cl	M-H, Random, 95% Cl
4.1.1 3 months							
Okun 2012	14	101	4	35	10.8%	1.21 [0.43, 3.44]	
Subtotal (95% CI)		101		35	10.8%	1.21 [0.43, 3.44]	
Total events	14		4				
Heterogeneity: Not applicable							
Test for overall effect: $Z = 0.36$ (P = 0.72)							
4.1.2 6 months							
Deuschl 2006	10	78	3	35	8.0%	1.50 [0.44, 5.10]	
Weaver 2009	49	134	15	134	33.4%	3.27 [1.93, 5.53]	
Subtotal (95% CI)		212		169	41.5%	2.71 [1.41, 5.21]	
Total events	59		18				
Heterogeneity: Tau ² = 0.07; Chi ² = 1.32, d	f = 1 (P = 0	0.25); I²	= 24%				
Test for overall effect: $Z = 3.00 (P = 0.003)$)						
4.1.3 12 months							
PDSURG full population (Williams 2010)	65	183	26	157	47.7%	2.14 [1.44, 3.21]	
Subtotal (95% CI)		183		157	47.7%	2.14 [1.44, 3.21]	
Total events	65		26				
Heterogeneity: Not applicable							
Test for overall effect: $Z = 3.72$ (P = 0.000)	2)						
Total (95% CI)		496		361	100.0%	2.26 [1.57, 3.23]	•
Total events	138		48				
Heterogeneity: $Tau^2 = 0.03$; $Chi^2 = 3.74$, d	f = 3 (P = 0	0.29); l²	= 20%				
Test for overall effect: Z = 4.43 (P < 0.000	01)						U.1 U.2 U.5 1 2 5 1 Eavours DBS Eavours medication
	df = 2 (D	- 0 44)	12 - 0%				Favous DBS Favous medication

DBS -v- medication alone: serious adverse events (proportion of participants experiencing ≥1) – forest plot

			DBS	Medication		Incidence rate ratio		Incidenc	e rate rat	io	
Study or Subgroup	log[Incidence rate ratio]	SE	Total	Total	Weight	IV, Random, 95% CI		IV, Rand	om, 95%	CI	
4.2.1 3 months					0	, ,					
Okun 2012	-0.00995	0.439155	101	35	20.6%	0.99 [0.42, 2.34]			+		
Subtotal (95% CI)			101	35	20.6%	0.99 [0.42, 2.34]					
Heterogeneity: Not applicable											
Test for overall effect: Z = 0.02 (P = 0.98)											
4.2.2 6 months											
Deuschl 2006	1.203973	0.658281	78	35	12.3%	3.33 [0.92, 12.11]			1	-	-
Weaver 2009	1.360231	0.254611	134	134	32.0%	3.90 [2.37, 6.42]					_
Subtotal (95% CI)			212	169	44.3%	3.82 [2.40, 6.08]					*
Heterogeneity: Tau ² = 0.00; Chi ² = 0.05, df	= 1 (P = 0.82); I ² = 0%										
Test for overall effect: Z = 5.64 (P < 0.0000)1)										
4.2.3 12 months											
PDSLIPG full population (Williams 2010)	1 197052	0 211805	183	157	35 1%	3 31 [2 10 5 01]				_	
Subtotal (95% CI)	1.137002	0.211035	183	157	35.1%	3.31 [2.19, 5.01]					
Heterogeneity: Not applicable										-	
Test for overall effect: $Z = 5.65$ (P < 0.0000)1)										
Total (95% CI)			496	361	100.0%	2.72 [1.60, 4.64]					
Heterogeneity: Tau ² = 0.16; Chi ² = 7.66, df	= 3 (P = 0.05); I ² = 61%							0.5	$\frac{1}{1}$	5	10
Test for overall effect: Z = 3.69 (P = 0.0002	2)						5.1 0.2 F	avours DBS	Favour	s medic:	ation
Test for subgroup differences: Chi ² = 7.61,	df = 2 (P = 0.02), $I^2 = 73.7\%$									2	

DBS -v- medication alone: serious adverse events (rate per patient-year) – forest plot

	DBS	5	Medica	tion		Risk Ratio		Risk Ratio	
Study or Subgroup	Events	Total	Events	Total	Weight	M-H, Random, 95% C		M-H, Random, 95% Cl	
4.4.1 3 months									
Okun 2012	9	101	0	35	15.3%	6.71 [0.40, 112.31]			
Subtotal (95% CI)		101		35	15.3%	6.71 [0.40, 112.31]			
Total events	9		0						
-leterogeneity: Not applicable									
Test for overall effect: $Z = 1.32$ (P = 0.19)									
4.4.2 6 months									
Deuschl 2006	1	78	1	35	15.9%	0.45 [0.03, 6.97]	-		
Neaver 2009	16	134	6	134	37.5%	2.67 [1.08, 6.61]			
3ubtotal (95% Cl)		212		169	53.3%	1.78 [0.41, 7.68]			
Total events	17		7						
Heterogeneity: Tau ² = 0.50; Chi ² = 1.46, df	^f = 1 (P = ().23); l²	= 32%						
Test for overall effect: $Z = 0.78$ (P = 0.44)									
4.4.3 12 months									
PDSURG full population (Williams 2010)	3	183	7	157	31.4%	0.37 [0.10, 1.40]			
Subtotal (95% CI)		183		157	31.4%	0.37 [0.10, 1.40]			
Total events	3		7						
leterogeneity: Not applicable									
Test for overall effect: $Z = 1.47$ (P = 0.14)									
Fotal (95% CI)		496		361	100.0%	1.24 [0.32, 4.83]			
Total events	29		14						
Heterogeneity: Tau² = 1.07; Chi² = 7.66, df	= 3 (P = 0	0.05); l²	= 61%				+		
Test for overall effect: $Z = 0.31 (P = 0.75)$							0.01	U.1 1 10 Eavours DBS Eavours modication	100 מר
Test for subgroup differences: $Chi^2 = 4.50$.	df = 2 (P	= 0.11)	. ² = 55.6	%					

DBS -v- medication alone: falls (proportion of participants experiencing ≥1) – forest plot

				Incidence rate ratio		Incidence	e rate ratio	
Study or Subgroup	log[Incidence rate ratio]	SE	Weight	IV, Random, 95% C		IV, Rand	om, 95% Cl	
4.5.1 3 months								
Okun 2012	1.884667	1.450953	13.0%	6.58 [0.38, 113.12]				
Subtotal (95% CI)			13.0%	6.58 [0.38, 113.12]				
Heterogeneity: Not applicable								
Test for overall effect: $Z = 1.30$ (P = 0.19)								
4.5.2 6 months								
Deuschl 2006	0	1.414214	13.5%	1.00 [0.06, 15.99]			+	-
Weaver 2009	0.939405	0.474858	41.5%	2.56 [1.01, 6.49]				
Subtotal (95% CI)			55.0%	2.33 [0.96, 5.62]				
Heterogeneity: Tau ² = 0.00; Chi ² = 0.40, df	= 1 (P = 0.53); I ² = 0%							
Test for overall effect: $Z = 1.88$ (P = 0.06)								
4.5.3 12 months								
PDSURG full population (Williams 2010)	-0.8473	0.690066	32.1%	0.43 [0.11, 1.66]			+-	
Subtotal (95% CI)			32.1%	0.43 [0.11, 1.66]				
Heterogeneity: Not applicable								
Test for overall effect: $Z = 1.23$ (P = 0.22)								
Total (95% CI)			100.0%	1.44 [0.45, 4.62]				
Heterogeneity: Tau ² = 0.63; Chi ² = 5.67, df	= 3 (P = 0.13); I ² = 47%				+		+	
Test for overall effect: Z = 0.61 (P = 0.54)					0.01	U.1 Eavours DBS	T 10	100 Jucation
Test for subgroup differences: Chi ² = 5.28,	df = 2 (P = 0.07), l ² = 62.1%	D					i avouis meu	ication

DBS -v- medication alone: falls (rate per patient-year) – forest plot

E.6.1.2 Symptom severity

Quality ass	essment					Number	of patients	Effect	
Number of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	DBS	Control	Mean difference (MD) (95% CI)	Quality
Hoehn and	Yahr scor	e (off medicatio	on) (lower is bette	er); 3–12 months					
3 ^{1,3,4}	RCT	No serious	No serious	Serious ⁵	No serious	261	215	-0.66 (-0.82 to -0.50)	MODERATE
Daily 'on' tir	me withou	t troublesome	dysinkesias (high	ner is better); 3–6	6 months				
2 ^{1,3}	RCT	No serious	Serious ⁸	Serious ⁵	No serious	275	229	3.66 (1.62 to 5.71)	LOW
Daily 'off' ti	me (lower	is better); 6-12	2 months						
2 ^{3,4}	RCT	No serious	No serious	Very serious ^{5,9}	No serious	169	185	-2.48 (-3.10 to -1.86)	LOW
UPDRS I (lo	wer is bet	ter); 3–12 mon	ths						
4 ^{1,2,3,4}	RCT	No serious	No serious	Serious ⁵	No serious	323	281	-0.29 (-0.60 to 0.02)	MODERATE
UPDRS II or	n (lower is	better); 3–12 r	nonths						
4 ^{1,2,3,10}	RCT	No serious	No serious ⁷	Serious ⁵	No serious	352	276	-2.98 (-4.50 to -1.46)	MODERATE
UPDRS III o	n (lower is	s better); 3–12	months						
4 ^{1,2,3,10}	RCT	No serious	Serious ⁸	Serious ⁵	No serious	331	280	-4.93 (-7.52 to -2.34)	LOW
UPDRS IV (lower is be	etter); 3–12 mo	nths						
3 ^{1,3,4}	RCT	No serious	Serious ⁸	Serious ⁵	No serious	243	204	-4.05 (-5.83 to -2.28)	LOW

¹ Okun 2012

² Deuschl 2006

³ Weaver 2009

⁴ PDSURG observed cases; ANCOVA model with baseline score, Hoehn and Yahr status (<3 -v- ≥3) and treatment allocation as covariates of final score used to estimate treatment effect in people with Hoehn and Yahr score ≥3 at baseline; calculated by guideline developers from patient-level data supplied by investigators (NB HY score ≥3 was a prespecified subgroup in the trial protocol and a randomisation stratification variable)

⁵ Some RCTs include a nontrivial proportion of participants with less advanced PD than the specified population for this question

⁶ At a 95% confidence level, data are only consistent with no meaningful effect

⁷ Some heterogeneity between 3-month and 6–12-month results; however direction of effect modification appears consistent and plausible, so not downgraded

⁸ *I*² greater than 40% with no obvious explanation for heterogeneity

⁹ PDSURG off time estimate approximated from answer to UPDRS Q39 (categorical proportion of waking day spent 'off')

¹⁰ PDSURG multiply imputed data; ANCOVA model with baseline score, Hoehn and Yahr status (<3 -v- ≥3) and treatment allocation as covariates of final score used to estimate treatment effect in people with Hoehn and Yahr score ≥3 at baseline; calculated by guideline developers from patient-level data supplied by investigators (NB HY score ≥3 was a prespecified subgroup in the trial protocol and a randomisation stratification variable)</p>

				Mean Difference		Mean L	Difference		
Study or Subgroup	Mean Difference	SE	Weight	IV, Fixed, 95% C	I	IV, Fixe	ed, 95% Cl		
3.1.1 3 months									
Okun 2012	-0.57	0.165784	24.6%	-0.57 [-0.89, -0.25]					
Subtotal (95% CI)			24.6%	-0.57 [-0.89, -0.25]					
Heterogeneity: Not applicable									
Test for overall effect: $Z = 3.44$ (P = 0.0006)									
3.1.2 6 months									
Weaver 2009	-0.7	0.112867	53.1%	-0.70 [-0.92, -0.48]					
Subtotal (95% CI)			53.1%	-0.70 [-0.92, -0.48]		\bullet			
Heterogeneity: Not applicable									
Test for overall effect: $Z = 6.20 (P < 0.00001)$									
3.1.3 12 months									
PDSURG (baseline HY≥3) ANCOVA; OC*	-0.655	0.174	22.3%	-0.66 [-1.00, -0.31]					
PDSURG (baseline HY≥3) OC†	-0.621	0.198637	0.0%	-0.62 [-1.01, -0.23]		-			
Subtotal (95% CI)			22.3%	-0.66 [-1.00, -0.31]					
Heterogeneity: Not applicable									
Test for overall effect: $Z = 3.76$ (P = 0.0002)									
Total (95% CI)			100.0%	-0.66 [-0.82, -0.50]		•			
Heterogeneity: $Chi^2 = 0.42$, df = 2 (P = 0.81);	l² = 0%				<u>ا</u>		+		
Test for overall effect: Z = 8.00 (P < 0.00001)					-2	-1 Favours DBS		I medication	2
Test for subgroup differences: Chi ² = 0.42, df	= 2 (P = 0.81), l ² =	0%				i avouis DBC		neuleation	

DBS -v- medication alone: Hoehn and Yahr score (off medication) – forest plot

	D	BS		Medication				Mean Difference	Mean Difference	
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Random, 95% CI	IV, Random, 95% CI	
3.2.1 3 months										
Okun 2012	4.3	4.2	101	1.8	4.2	34	44.7%	2.50 [0.87, 4.13]		
Subtotal (95% CI)			101			34	44.7%	2.50 [0.87, 4.13]		
Heterogeneity: Not ap	plicable									
Test for overall effect:	Z = 3.00	(P =	0.003)							
3.2.2 6 months		_								
Deuschl 2006	4.4	0	53	-0.5	0	61		Not estimable	_	
Weaver 2009	4.6	4.4	121	0	2.9	134	55.3%	4.60 [3.67, 5.53]		
Subtotal (95% CI)			174			195	55.3%	4.60 [3.67, 5.53]		
Heterogeneity: Not ap	plicable									
Test for overall effect:	Z = 9.75	(P <	0.00001	1)						
3.2.3 12 months										
Subtotal (95% CI)			0			0		Not estimable		
Heterogeneity: Not ap	plicable									
Test for overall effect:	Not appli	cable	;							
			075			220	100.0%	2 66 14 62 5 741		
10tal (95% CI)			215			229	100.0%	J.00 [1.02, J./1]		
Heterogeneity: Tau ² = 1.75; Chi ² = 4.81, df = 1 (P = 0.03); $I^2 = 79\%$								-4 -2 0 2 4		
Test for overall effect: Z = 3.51 (P = 0.0005)									Favours medication Favours DBS	
Test for subgroup diffe	erences: C	Chi² =	= 4.81, c	lf = 1 (I	⊃ = 0.	.03), l² :	= 79.2%			

DBS -v- medication alone: mean daily 'on' time without troublesome dysinkesias – forest plot
						wiea		ice	
Study or Subgroup N	lean Difference	SE	Weight	IV, Fixed, 95% CI		IV,	Fixed, 95%	% CI	
3.3.1 3 months									
Subtotal (95% CI)				Not estimable					
Heterogeneity: Not applicable									
Test for overall effect: Not applicable									
3.3.2 6 months									
Deuschl 2006	-4.2	0		Not estimable					
Weaver 2009	-2.4	0.394635	63.5%	-2.40 [-3.17, -1.63]	-				
Subtotal (95% CI)			63.5%	-2.40 [-3.17, -1.63]					
Heterogeneity: Not applicable									
Test for overall effect: $Z = 6.08$ (P < 0.00001)									
3.3.3 12 months									
PDSURG (baseline HY≥3) ANCOVA; MI*	-2.62	0.52	36.5%	-2.62 [-3.64, -1.60]		-			
PDSURG (baseline HY≥3) OC†	-2.971	0.598179	0.0%	-2.97 [-4.14, -1.80]		-			
Subtotal (95% CI)			36.5%	-2.62 [-3.64, -1.60]					
Heterogeneity: Not applicable									
Test for overall effect: $Z = 5.04$ (P < 0.00001)									
Total (95% CI)			100.0%	-2.48 [-3.10, -1.86]		•			
Heterogeneity: $Chi^2 = 0.11$, $df = 1$ (P = 0.74); I^2	= 0%			-	<u> </u>				<u> </u>
Test for overall effect: $Z = 7.89 (P < 0.00001)$					-4	-2		2	4
Test for subgroup differences: $Chi^2 = 0.11$, df =	1 (P = 0.74), I ² =	0%				Favours	JBS Fav	ours mea	Ication

DBS -v- medication alone: mean daily 'off' time – forest plot

Otrada an Orahamaan	Differences	05	14/- :			NIC N		,	
Study or Subgroup M	ean Difference	SE	weight	IV, Fixed, 95% C		IV	Fixed, 95% C	71	
3.4.1 3 months									
Okun 2012	-0.01	0.343507	20.9%	-0.01 [-0.68, 0.66]				-	
Subtotal (95% CI)			20.9%	-0.01 [-0.68, 0.66]				-	
Heterogeneity: Not applicable									
Test for overall effect: Z = 0.03 (P = 0.98)									
3.4.2 6 months									
Deuschl 2006	-0.58	0.409376	14.8%	-0.58 [-1.38, 0.22]					
Weaver 2009	-0.3	0.253336	38.5%	-0.30 [-0.80, 0.20]					
Subtotal (95% CI)			53.3%	-0.38 [-0.80, 0.04]					
Heterogeneity: Chi ² = 0.34, df = 1 (P = 0.56); l ² =	= 0%								
Test for overall effect: Z = 1.75 (P = 0.08)									
3.4.3 12 months									
PDSURG (baseline HY≥3) ANCOVA; OC*	-0.32744	0.309642	25.8%	-0.33 [-0.93, 0.28]					
PDSURG (baseline HY≥3) OC†	-0.246	0.391106	0.0%	-0.25 [-1.01, 0.52]					
Subtotal (95% CI)			25.8%	-0.33 [-0.93, 0.28]					
Heterogeneity: Not applicable									
Test for overall effect: $Z = 1.06 (P = 0.29)$									
Total (95% CI)			100.0%	-0.29 [-0.60, 0.02]		-			
Heterogeneity: Chi ² = 1.18, df = 3 (P = 0.76); l ² =	= 0%				⊢	<u> </u>	<u> </u>		
Test for overall effect: Z = 1.83 (P = 0.07)					-2	-1 Eavoura		1 re modicatio	2
Test for subgroup differences: $Chi^2 = 0.84$, df = 2	2 (P = 0.66), I ² =	0%				Favours	DBS Favour	s medicatio	11

DBS -v- medication alone: UPDRS I - forest plot

Study or Subaroup	Mean Difference	SE	Weiaht	IV. Random. 95% C	1	IV. Ran	dom. 95% Cl	
3.5.1 3 months				, ,		,		
Okun 2012	-0.91	1 285723	19 1%	-0 91 [-3 43 1 61]				
Subtotal (95% CI)			19.1%	-0.91 [-3.43, 1.61]				
Heterogeneity: Not applicable								
Test for overall effect: $Z = 0.71$ (P = 0.48)								
3.5.2 6 months								
Deuschl 2006	-2.6	0.880574	26.5%	-2.60 [-4.33, -0.87]			-	
Weaver 2009	-4.6	0.643937	31.6%	-4.60 [-5.86, -3.34]				
Subtotal (95% CI)			58.1%	-3.69 [-5.64, -1.74]				
Heterogeneity: Tau² = 1.40; Chi² = 3.36, d	f = 1 (P = 0.07); l ² = 7	0%						
Test for overall effect: Z = 3.71 (P = 0.000	2)							
3.5.3 12 months								
PDSURG (baseline HY≥3) ANCOVA; MI*	-2.93	1.07	22.8%	-2.93 [-5.03, -0.83]			-	
PDSURG (baseline HY≥3) OC†	-3.762	1.23112	0.0%	-3.76 [-6.17, -1.35]				
Subtotal (95% CI)			22.8%	-2.93 [-5.03, -0.83]				
Heterogeneity: Not applicable								
Test for overall effect: $Z = 2.74$ (P = 0.006)							
Total (95% CI)			100.0%	-2.98 [-4.50, -1.46]		•		
Heterogeneity: Tau² = 1.49; Chi² = 8.24, d	f = 3 (P = 0.04); I ² = 6	4%			10	— 		10
Test for overall effect: Z = 3.85 (P = 0.000	1)				-10	-ə Favours DR	S Favours med	10 lication
Test for subgroup differences: Chi ² = 2.97	, df = 2 (P = 0.23), l² =	32.6%						loadon

DBS -v- medication alone: UPDRS II (on) - forest plot

Study or Subaroup	Mean Difference	SE	Weight	IV Random 95% C	I	IV R	andom 95% (21	
3.6.1 3 months				,		,			
Okun 2012 Subtotal (95% CI)	-7.38	1.434864	25.2% 25.2%	-7.38 [-10.19, -4.57] -7.38 [-10.19, -4.57]		•			
Heterogeneity: Not applicable									
Test for overall effect: $Z = 5.14$ (P < 0.0000	1)								
3.6.2 6 months									
Deuschl 2006	-4.4	1.499651	24.5%	-4.40 [-7.34, -1.46]					
Weaver 2009	-2	1.136584	28.3%	-2.00 [-4.23, 0.23]		-			
Subtotal (95% CI)			52.9%	-3.00 [-5.32, -0.68]					
Heterogeneity: Tau ² = 1.11; Chi ² = 1.63, df	= 1 (P = 0.20); l ² = 3	9%							
Test for overall effect: $Z = 2.54$ (P = 0.01)									
3.6.3 12 months									
PDSURG (baseline HY≥3) ANCOVA; MI*	-6.496	1.761	21.9%	-6.50 [-9.95, -3.04]			-		
PDSURG (baseline HY≥3) OC†	-9.493	2.214053	0.0%	-9.49 [-13.83, -5.15]					
Subtotal (95% CI)			21.9%	-6.50 [-9.95, -3.04]			•		
Heterogeneity: Not applicable									
Test for overall effect: Z = 3.69 (P = 0.0002)								
Total (95% CI)			100.0%	-4.93 [-7.52, -2.34]		-	▶		
Heterogeneity: Tau ² = 4.88; Chi ² = 10.15, d	f = 3 (P = 0.02); l ² =	70%			H				<u> </u>
Test for overall effect: Z = 3.73 (P = 0.0002)				-20	-10 Favoure f		10 medicatic	20 20
Test for subgroup differences: $Chi^2 = 6.30$,	df = 2 (P = 0.04), l ² =	68.2%						medicatio	///

DBS -v- medication alone: UPDRS III (on) - forest plot

Other day and Outhermouth		65		Mean Difference		Mean Diffe	ence	
Study or Subgroup	Mean Difference	SE	weight	IV, Random, 95% C	1	IV, Random,	95% CI	
3.7.1 3 months						_		
Okun 2012	-3.4	0.612372	32.9%	-3.40 [-4.60, -2.20]				
Subtotal (95% CI)			32.9%	-3.40 [-4.60, -2.20]				
Heterogeneity: Not applicable								
Test for overall effect: $Z = 5.55$ (P < 0.00001)							
3.7.2 6 months								
Weaver 2009	-2.9	0.434124	35.5%	-2.90 [-3.75, -2.05]				
Subtotal (95% CI)			35.5%	-2.90 [-3.75, -2.05]		•		
Heterogeneity: Not applicable								
Test for overall effect: $Z = 6.68$ (P < 0.00001)							
3.7.3 12 months								
PDSURG (baseline HY≥3) ANCOVA; OC*	-6.03297	0.692522	31.6%	-6.03 [-7.39, -4.68]	_	-		
PDSURG (baseline HY≥3) OC†	-5.495	0.833824	0.0%	-5.50 [-7.13, -3.86]		•		
Subtotal (95% CI)			31.6%	-6.03 [-7.39, -4.68]				
Heterogeneity: Not applicable								
Test for overall effect: Z = 8.71 (P < 0.00001)							
Total (95% CI)			100.0%	-4.05 [-5.83, -2.28]				
Heterogeneity: Tau ² = 2.12; Chi ² = 14.98, df	= 2 (P = 0.0006); l ² :	= 87%				<u> </u>	<u> </u>	
Test for overall effect: Z = 4.47 (P < 0.00001)				-10		5 wours medicativ	10
Test for subgroup differences: Chi ² = 14.98,	df = 2 (P = 0.0006),	l² = 86.7%						

DBS -v- medication alone: UPDRS IV – forest plot

E.6.1.3 Neuropsychological outcomes

Quality ass	essment					Number	r of patients		
Number	Decign	Bick of bicc	Inconsistancy	Indiractocco	Improvision	DPS	Control	Effect (0.5% CI)	Quality
of studies	Design	RISK UI DIAS	inconsistency	indirectiless					Quality
Cognitive f	unction (d	lifferent measu	ires pooled [stan	dardised mean	difference]) (n	igner is b	etter); 6–12 r	nonths	
3 ^{2,3,4}	RCT	No serious	Serious ⁵	Serious ⁶	Serious ⁷	310	334	SMD = -0.16 (-0.34 to 0.03)	VERY LOW
Semantic f	luency (hi	gher is better);	3–12 months						
4 ^{1,2,3,4}	RCT	No serious	No serious	Serious ⁶	Serious ⁷	324	271	SMD = -0.34 (-0.50 to -0.17)	LOW
Phonemic	fluency (h	igher is better)	; 6–12 months						
3 ^{2,3,4}	RCT	No serious	No serious	Serious ⁶	No serious	222	235	SMD = -0.52 (-0.71 to -0.33)	MODERATE
Depression	n (differen	t measures po	oled [standardise	ed mean differe	nce]) (lower is	better); 3	–6 months		
3 ^{1,2,3}	RCT	No serious	Serious ⁵	Serious ⁶	Very serious ⁸	274	233	SMD = -0.17 (-0.58 to 0.25)	VERY LOW
¹ Okun 201 ² Deuschl 2 ³ Weaver 2 ⁴ Williams	2 2006 (semai 2009 2010 (main	ntic fluency and p	honemic fluency rep tion fall participants	ported for a subgro	oup of participants	s in Witt 20 up data av	09) ailable for this o	utcome)	

⁵ I² greater than 40% with no obvious explanation for heterogeneity
 ⁶ Some RCTs include a nontrivial proportion of participants with less advanced PD than the specified population for this question
 ⁷ At a 95% confidence level, data are consistent with appreciable harm and no meaningful effect
 ⁸ At a 95% confidence level, data are consistent with appreciable benefit, appreciable harm and no meaningful effect

	l	DBS		Med	icatio	n	5	Std. Mean Difference	Std. Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Fixed, 95% CI	IV, Fixed, 95% Cl
3.9.1 3 months									
Subtotal (95% CI)			0			0		Not estimable	
Heterogeneity: Not applicable									
Test for overall effect: Not applicable									
3.9.2 6 months									
Deuschl 2006	-2	4.9	68	-0.5	4	67	20.8%	-0.33 [-0.67, 0.01]	
Weaver 2009	-0.2	6.1	121	0.9	4.7	134	39.6%	-0.20 [-0.45, 0.04]	
Subtotal (95% CI)			189			201	60.4%	-0.25 [-0.45, -0.05]	\bullet
Heterogeneity: $Chi^2 = 0.37$, df = 1 (P = 0.54)	4); I² = 0	%							
Test for overall effect: Z = 2.43 (P = 0.01)									
3.9.3 12 months									
PDSURG full population (Williams 2010)	-0.4	3.5	121	-0.4	2.9	133	39.6%	0.00 [-0.25, 0.25]	
Subtotal (95% CI)			121			133	39.6%	0.00 [-0.25, 0.25]	\bullet
Heterogeneity: Not applicable									
Test for overall effect: Z = 0.00 (P = 1.00)									
Total (95% CI)			310			334	100.0%	-0.15 [-0.30, 0.01]	◆
Heterogeneity: Chi ² = 2.72, df = 2 (P = 0.26	6); I² = 2	6%						-	
Test for overall effect: Z = 1.89 (P = 0.06)									- I -U.5 U U.5 1 Favours medication Favours DBS
Test for subgroup differences: Chi ² = 2.35,	df = 1 (P = 0	.13), l²	= 57.4%)				

DBS -v- medication alone: cognitive function (different measures pooled [standardised mean difference]) – forest plot

		DBS		N	ledication		:	Std. Mean Difference	Std. Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Fixed, 95% CI	IV, Fixed, 95% Cl
3.10.1 3 months									
Okun 2012	-1.9	2.627233	101	-1.52	2.627233	35	18.8%	-0.14 [-0.53, 0.24]	
Subtotal (95% CI)			101			35	18.8%	-0.14 [-0.53, 0.24]	
Heterogeneity: Not applicable									
Test for overall effect: Z = 0.73 (P = 0.46)									
3.10.2 6 months									
Deuschl 2006	-6.1	11.6	31	0.3	10.3	31	10.8%	-0.58 [-1.08, -0.07]	
Weaver 2009	-4.7	11.1115	121	-2	11.1196	134	45.7%	-0.24 [-0.49, 0.00]	
Subtotal (95% CI)			152			165	56.5%	-0.31 [-0.53, -0.08]	
Heterogeneity: $Chi^2 = 1.34$, df = 1 (P = 0.28)	5); l² = 2	5%							
Test for overall effect: Z = 2.70 (P = 0.007)									
3.10.3 12 months									
PDSURG full population (Williams 2010)	-4.5	7.8	71	-0.2	7.7	71	24.7%	-0.55 [-0.89, -0.22]	
Subtotal (95% CI)			71			71	24.7%	-0.55 [-0.89, -0.22]	
Heterogeneity: Not applicable									
Test for overall effect: Z = 3.23 (P = 0.001)									
Total (95% CI)			324			271	100.0%	-0.34 [-0.50, -0.17]	◆
Heterogeneity: Chi ² = 3.96, df = 3 (P = 0.23	7); l² = 2	4%							+ + +
Test for overall effect: Z = 3.95 (P < 0.0001)								-1 -0.5 0 0.5
Test for subgroup differences: Chi ² = 2.62,	df = 2 (F	⊃ = 0.27), l²	= 23.7	%					Favours medication Favours DBS

DBS -v- medication alone: semantic fluency – forest plot

		DBS		M	edication	1	s	Std. Mean Difference	Std. Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Fixed, 95% CI	IV, Fixed, 95% Cl
3.11.1 3 months									
Subtotal (95% CI)			0			0		Not estimable	
Heterogeneity: Not applicable									
Test for overall effect: Not applicable									
3.11.2 6 months									
Deuschl 2006	-1.9	8.1	31	-0.5	6	31	14.0%	-0.19 [-0.69, 0.31]	•
Weaver 2009	-3.5	8.3336	121	1.1	8.7786	134	55.8%	-0.54 [-0.79, -0.28]	
Subtotal (95% CI)			152			165	69 .8%	-0.47 [-0.69, -0.24]	
Heterogeneity: Chi ² = 1.43, df = 1 (P = 0.23	3); I² = 3	0%							
Test for overall effect: Z = 4.09 (P < 0.000	1)								
244242 months									
3:11.3 12 IIIOIIUIS									
PDSURG full population (Williams 2010)	-6.5	9.4	70	-0.6	8.7	70	30.2%	-0.65 [-0.99, -0.31]	
Subtotal (95% CI)			70			70	30.2%	-0.65 [-0.99, -0.31]	
Heterogeneity: Not applicable									
Test for overall effect: Z = 3.73 (P = 0.0002	2)								
Total (95% CI)			222			235	100.0%	-0.52 [-0.71, -0.33]	
Heterogeneity: $Chi^2 = 2.20$ df = 2.(P = 0.2)	3)· 12 - 0	0/_					,0		-+ $-+$ $+$ $+$ $+$
Therefore every left of $Z = 5.47$ (D $\neq 0.000$)	57, I – 9 243	/0							-1 -0.5 0 0.5
Test for overall effect: $Z = 5.47$ (P < 0.0000	יי א'								Favours medication Favours DBS
l est for subgroup differences: Chi ² = 0.76.	df = 1 (F	v = 0.38	$1^2 = 0$	6					

DBS -v- medication alone: phonemic fluency – forest plot

	Fav	ours DE	S	Me	dicatior	1	5	Std. Mean Difference	Std. Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Random, 95% CI	IV, Random, 95% CI
3.12.1 3 months									
Okun 2012	-9.14	12.1	88	-1.8	12.1	30	29.6%	-0.60 [-1.02, -0.18]	
Subtotal (95% CI)			88			30	29.6%	-0.60 [-1.02, -0.18]	
Heterogeneity: Not ap	plicable								
Test for overall effect:	Z = 2.80) (P = 0.0	05)						
3.12.2 6 months									
Deuschl 2006	-0.3	7.2	65	0.6	6.3	69	33.2%	-0.13 [-0.47, 0.21]	
Weaver 2009	-0.4	7.2225	121	-1.5	7.0229	134	37.1%	0.15 [-0.09, 0.40]	
Subtotal (95% CI)			186			203	70.4%	0.04 [-0.24, 0.31]	\bullet
Heterogeneity: Tau ² =	0.02; Cł	ni² = 1.80	, df = 1	(P = 0.	18); I² = 4	4%			
Test for overall effect:	Z = 0.25	6 (P = 0.8	0)						
3.12.3 12 months									
Subtotal (95% CI)			0			0		Not estimable	
Heterogeneity: Not ap	plicable								
Test for overall effect:	Not app	licable							
Total (95% CI)			274			233	100.0%	-0.17 [-0.58, 0.25]	
Heterogeneity: Tau ² =	0.10; Cł	1i² = 9.44	, df = 2	(P = 0.	009); I² =	79%			
Test for overall effect:	Z = 0.78	6 (P = 0.4	3)	-	-				-1 -0.5 0 0.5 1
Test for subgroup diffe	rancas.	$Chi^2 = 6$	15 df =	= 1 (P =	0 01) l ²	= 83 70	10		Favours DBS Favours medication

DBS -v- medication alone: depression (different measures pooled [standardised mean difference]) – forest plot

E.6.1.4 Health related quality of life – patient

Quality ass	essment					Number	of patients	Effect	1
Number of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	DBS	Control	Mean difference (MD) (95% CI)	Quality
EQ-5D (higl	her is bett	er); 12 months							
1 ⁴	RCT	No serious	No serious	No serious	No serious	50	50	0.123 (0.022 to 0.225)	HIGH
PDQ-39 (lov	wer is bett	er); 6–12 mont	hs						
3 ^{2,3,4}	RCT	No serious	No serious	Serious ⁵	No serious	243	258	-8.28 (-10.27 to -6.30)	MODERATE
¹ Okun 201 ² Deuschl 2	2 2006								

³ Weaver 2009

⁴ PDSURG multiply imputed data; ANCOVA model with baseline score, Hoehn and Yahr status (<3 -v- ≥3) and treatment allocation as covariates of final score used to estimate treatment effect in people with Hoehn and Yahr score ≥3 at baseline; calculated by guideline developers from patient-level data supplied by investigators (NB HY score ≥3 was a prespecified subgroup in the trial protocol and a randomisation stratification variable)

⁵ Some RCTs include a nontrivial proportion of participants with less advanced PD than the specified population for this question

		-		wean Difference		Mean D	interence	
Study or Subgroup N	lean Difference	SE	Weight	IV, Fixed, 95% C		IV, Fixe	ed, 95% Cl	
3.12.1 3 months								
Subtotal (95% CI)				Not estimable				
Heterogeneity: Not applicable								
Test for overall effect: Not applicable								
3.12.2 6 months								
Deuschl 2006	-9.7	2.2398	20.5%	-9.70 [-14.09, -5.31]	-			
Weaver 2009	-8.1	1.2755	63.1%	-8.10 [-10.60, -5.60]		-		
Subtotal (95% CI)			83.5%	-8.49 [-10.66, -6.32]		•		
Heterogeneity: $Chi^2 = 0.39$, df = 1 (P = 0.53); I ²	= 0%							
Test for overall effect: $Z = 7.66$ (P < 0.00001)								
3.12.3 12 months								
PDSURG (baseline HY≥3) ANCOVA; MI*	-7.219	2.495	16.5%	-7.22 [-12.11, -2.33]				
PDSURG (baseline HY≥3) OC†	-7.35	2.7603	0.0%	-7.35 [-12.76, -1.94]				
Subtotal (95% CI)			16.5%	-7.22 [-12.11, -2.33]				
Heterogeneity: Not applicable								
Test for overall effect: Z = 2.89 (P = 0.004)								
Total (95% CI)			100.0%	-8.28 [-10.27, -6.30]		•		
Heterogeneity: $Chi^2 = 0.60$, $df = 2$ (P = 0.74); I^2	= 0%					10	+ + 0 10	
Test for overall effect: Z = 8.18 (P < 0.00001)					-20	- IU Favours DBS	U 10 Favours medicati	20 on
Test for subgroup differences: Chi ² = 0.22, df =	1 (P = 0.64), I ² =	0%					i avours medicali	

DBS -v- medication alone: PDQ-39 – forest plot

E.6.1.5 Medication load

Quality ass	essment					Numbe	r of patients	Effect	
Number of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	DBS	Control	Mean difference (MD) (95% CI)	Quality
Daily dosag	ge of anti-l	Parkinson's me	dication (levodor	oa mg equivalent	t) (lower is bette	er); 3–6 n	nonths		
3 ^{1,2,3}	RCT	No serious	No serious	Serious ⁵	No serious	293	240	-381 (-468 to -295)	MODERATE
¹ Okun 201 ² Deuschl 2 ³ Weaver 2	2 2006								

³ Weaver 2009
 ⁴ Some RCTs include a nontrivial proportion of participants with less advanced PD than the specified population for this question

		DBS		N	ledication			Mean Difference	Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Fixed, 95% CI	IV, Fixed, 95% Cl
3.13.1 3 months									
Okun 2012	-492	437.005	101	-131	437.005	35	26.5%	-361.00 [-529.00, -193.00]	
Subtotal (95% CI)			101			35	26.5%	-361.00 [-529.00, -193.00]	
Heterogeneity: Not ap	plicable								
Test for overall effect:	Z = 4.21	(P < 0.000	1)						
3.13.2 6 months									
Deuschl 2006	-593	548	71	-95	390	71	30.6%	-498.00 [-654.45, -341.55]	_
Weaver 2009	-296	666.6901	121	15	339.4397	134	42.9%	-311.00 [-442.96, -179.04]	
Subtotal (95% CI)			192			205	73.5%	-388.74 [-489.61, -287.86]	\bullet
Heterogeneity: Chi ² =	3.21, df	= 1 (P = 0.0	07); l² =	69%					
Test for overall effect:	Z = 7.55	5 (P < 0.000	01)						
3.13.3 12 months									
Subtotal (95% CI)			0			0		Not estimable	
Heterogeneity: Not ap	plicable								
Test for overall effect:	Not app	licable							
Total (95% CI)			293			240	100.0%	-381.39 [-467.87, -294.91]	◆
Heterogeneity: Chi ² =	3.28, df	= 2 (P = 0.1	9); l ² =	39%				-	
Test for overall effect:	Z = 8.64	(P < 0.000	01)						-500 -250 0 250 500
Test for subgroup diffe	erences:	Chi ² = 0.08	, df = 1	(P = 0.1)	78), l² = 0%				Favours DBS Favours medication

DBS -v- medication alone: change in mean daily dose of anti-Parkinson's medication (levodopa mg equivalent) – forest plot

Levodopa-carbidopa intestinal gel compared with best medical treatment for advanced Parkinson's disease E.6.2

E.6.2.1 Adverse events

		Quality asse	essment			No. of no. of	events / patients				
No. of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	LCIG	Placebo- LCIG	Effect (95% CI)	Quality		
Serious adverse events (probability of experiencing ≥1)											
Olanow et. al. (2014)	RCT	High ¹	NA ²	Serious ³	Serious ⁴	5/37	7/34	RR = 0.66 (0.23 to 1.87)	VERY LOW		
Any adverse events	s (probabi	ility of experie	encing ≥1)								
Olanow et. al. (2014)	RCT	High ¹	NA ²	Serious ³	None	35/37	34/34	RR = 0.95 (0.86 to 1.04)	LOW		
Device complication	ns (proba	bility of expe	riencing ≥1)								
Olanow et. al. (2014)	RCT	High ¹	NA ²	Serious ³	Serious ⁴	34/37	29/34	RR = 1.08 (0.91 to 1.28)	VERY LOW		
Falls (probability of	experien	cing ≥1)									
Olanow et. al. (2014)	RCT	High ¹	NA ²	Serious ³	Serious ⁴	4/37	4/34	RR = 0.92 (0.25 to 3.39)	VERY LOW		

¹ High risk of bias, due to device implantation in both trial arms
² NA: Not applicable as only 1 study contributed to this analysis
³ Serious indirectness, due to device implantation in both trial arms

⁴ At a 95% confidence level, data are consistent with appreciable harm, appreciable benefit and no difference

E.6.2.2 Symptom severity

Quality assessment						Number patients	of G	Effect	
Number of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	LCIG	Control	Mean difference (MD) (95% Cl)	Quality
On time without dys	kinesias (hrs, increase i	s good)						
Olanow et. al. (2014)	RCT	Low ¹	NA ²	None ³	None	35	31	2.28 (0.4 to 4.09)	HIGH
Off time per day (hrs	, reductio	n is good)							
Olanow et. al. (2014)	RCT	Low ¹	NA ²	None ³	None	35	31	-1.91 (-3.03 to -0.79)	HIGH
UPDRS II (on) (lower	is better)								
Olanow et. al. (2014)	RCT	Low ¹	NA ²	None ³	None	35	31	-3.00 (-5.16 to -0.84)	HIGH
UPDRS III (on) (lower	r is better)							
Olanow et. al. (2014)	RCT	Low ¹	NA ²	None ³	Serious ⁴	35	31	1.40 (-2.72 to 5.52)	MODERATE
Clinical global impre	ssion of c	hange score (lower is better)						
Olanow et. al. (2014)	RCT	Low ¹	NA ²	None ³	None	35	31	-0.7 (-1.4 to -0.1)	HIGH
1 low risk of hige as asse	seed by NI	CE RCT quality of	hacklist						

¹ Low risk of bias, as assessed by NICE RCT quality checklist ² NA: Not applicable as only 1 study contributed to this analysis

³ No serious indirectness; population was as described in review protocol
 ⁴ At a 95% confidence level, data are consistent with appreciable harm, appreciable benefit and no difference

E.6.2.3 Health-related quality of life - patient

Quality assessment						Number patients	r of S	Effect	
Number of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	LCIG	Control	Mean difference (MD) (95% Cl)	Quality
Generic health-relate	ed quality	of life: EQ-5D							
Olanow et. al. (2014)	RCT	Low ¹	NA ²	None ³	Serious ⁴	35	31	0.07 (-0.01 to 0.15)	MODERATE
Parkinson's disease	-related q	juality of life: F	PDQ 39						
Olanow et. al. (2014)	RCT	Low ¹	NA ²	None ³	None	35	31	-7.00 (-12.49 to - 1.51)	HIGH
1 I am state of black as as as	and a share but	OF DOT availt	a la a lulia t						

Low risk of bias, as assessed by NICE RCT quality checklist

² NA: Not applicable as only one study contributed to this analysis

³ No serious indirectness; population was as described in review protocol
 ⁴ At a 95% confidence level, data are consistent with appreciable benefit and no difference

E.6.2.4 Health-related quality of life – carer

Quality assessment				Quality assessment						
Number of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	LCIG	Control	Mean difference (MD) (95% CI)	Quality	
Zarit carer burden in	terview									
Olanow et. al. (2014)	RCT	Low ¹	NA ²	None ³	Serious ⁴	35	31	-4.5 (-10.58 to 1.58)	MODERATE	

¹ Low risk of bias, as assessed by NICE RCT quality checklist ² NA: Not applicable as only one study contributed to this analysis

³ No serious indirectness; population was as described in review protocol

⁴ At a 95% confidence level, data are consistent with appreciable benefit and no difference

E.6.2.5 Medication load

Quality assessment						Number patients	r of S	Effect	
Number of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	LCIG	Control	Mean difference (MD) (95% Cl)	Quality
Levodopa daily dosa	age (mg)								
Olanow et. al. (2014)	RCT	Low ¹	NA ²	None ³	Serious ⁴	35	31	-158.0 (-324.5 to 8.5)	MODERATE
Number of studies Levodopa daily dosa Olanow et. al. (2014)	Design age (mg) RCT	bias	Inconsistency	Indirectness None ³	Imprecision Serious ⁴	LCIG 35	Control 31	(MD) (95% CI) -158.0 (-324.5 to 8.5)	Quali MOD

¹ Low risk of bias, as assessed by NICE RCT quality checklist ² NA: Not applicable as only one study contributed to this analysis

³ No serious indirectness; population was as described in review protocol

⁴ At a 95% confidence level, data are consistent with appreciable benefit and no difference

E.6.3 Indirect comparison of DBS and LCIG

E.6.3.1 Symptom severity

			Pairwise	Direct evidence		Indirect evidence		
Comparison	Studies	Timepoint	data	Effect measure (95%Cl)	Quality of evidence	Effect measure (95%Cl)	Quality of evidence	
UPDRS II (lower is better)								
DBS (n=45) -v- BMT (n=47)	PDSURG (HY≥3) ⁶	52wk	E.6.1.2	-2.92 (-5.02 to -0.82)	HIGH	-	-	
LCIG (n=35) -v- BMT (n=31)	Olanow et al. (2014)	12wk	E.6.2.2	-3.00 (-5.16 to -0.84)	HIGH	-	-	
DBS -v- LCIG	-	52wk ¹	_	-	-	0.08 (-3.14 to 3.29)	LOW ^{2,3}	
UPDRS III (lower is better)								
DBS (n=40) -v- BMT (n=38)	PDSURG (HY≥3) ⁶	52wk	E.6.1.2	-6.48 (-9.93 to -3.03)	HIGH	-	-	
LCIG (n=35) -v- BMT (n=31)	Olanow et al. (2014)	12wk	E.6.2.2	1.40 (-2.72 to 5.52)	MODERATE ⁴	-	-	
DBS -v- LCIG	-	52wk ¹	_	-	-	-7.88 (-13.63 to -2.14)	MODERATE ²	
Off time (lower is better)								
DBS (n=48) -v- BMT (n=51)	PDSURG (HY≥3) ⁶	52wk	E.6.1.2	-2.62 (-3.65 to -1.60)	MODERATE ⁵	-	-	
LCIG (n=35) -v- BMT (n=31)	Olanow et al. (2014)	12wk	E.6.2.2	-1.91 (-3.03 to -0.79)	HIGH	-	-	
DBS -v- LCIG	-	52wk ¹	_	_	-	-0.71 (-2.29, 0.87)	VERY LOW ^{2,3,5}	

¹ Incorporating increased uncertainty for LCIG -v- BMT due to unknown 'drift' from 12wk to 52wk timepoints (parameterised using Femandez et al. 2015)

² Downgraded for indirectness (12wk estimate used to estimate 52wk effects)

³ Downgraded for imprecision (at a 95% confidence level, data are consistent with appreciable benefit with DBS, appreciable benefit with LCIG and no meaningful difference)

⁴ Downgraded for imprecision

⁵ Downgraded for indirectness (off time estimate approximated from answer to UPDRS Q39 [categorical proportion of waking day spent 'off'])

⁶ PDSURG multiply imputed data; ANCOVA model with baseline score, Hoehn and Yahr status (<3 -v- ≥3) and treatment allocation as covariates of final score used to estimate treatment effect in people with Hoehn and Yahr score ≥3 at baseline; calculated by guideline developers from patient-level data supplied by investigators (NB HY score ≥3 was a prespecified subgroup in the trial protocol and a randomisation stratification variable)</p>

E.6.3.2 Health-related quality of life – patient

			Pairwise	Direct evidence		Indirect evidence		
Comparison	Studies	Timepoint	data	Effect measure (95%Cl)	Quality of evidence	Effect measure (95%Cl)	Quality of evidence	
EQ-5D (higher is better)								
DBS (n=50) -v- BMT (n=50)	PDSURG (HY≥3) ⁵	52wk	E.6.1.4	0.12 (0.02 to 0.22)	HIGH	_	_	
LCIG (n=35) -v- BMT (n=31)	Olanow et al. (2014)	12wk	E.6.2.3	0.07 (-0.01 to 0.15)	MODERATE ⁴	-	_	
DBS -v- LCIG	-	52wk ¹	-	-	-	0.05 (-0.08 to 0.19)	LOW ^{2,3}	
PDQ-39 (lower is better)								
DBS (n=51) -v- BMT (n=51)	PDSURG (HY≥3) ⁵	52wk	E.6.1.4	-7.21 (-12.10 to -2.32)	HIGH	_	_	
LCIG (n=35) -v- BMT (n=31)	Olanow et al. (2014)	12wk	E.6.2.3	-7.00 (-12.49 to -1.51)	HIGH	-	-	
DBS -v- LCIG	-	52wk ¹	_	_	_	-0.21 (-7.92 to 7.50)	LOW ^{2,3}	

¹ Incorporating increased uncertainty for LCIG -v- BMT due to unknown 'drift' from 12wk to 52wk timepoints (parameterised using Femandez et al. 2015)

² Downgraded for indirectness (12wk estimate used to estimate 52wk effects)

³ Downgraded for imprecision (at a 95% confidence level, data are consistent with appreciable benefit with DBS, appreciable benefit with LCIG and no meaningful difference)

⁴ Downgraded for imprecision

⁵ PDSURG multiply imputed data; ANCOVA model with baseline score, Hoehn and Yahr status (<3 -v- ≥3) and treatment allocation as covariates of final score used to estimate treatment effect in people with Hoehn and Yahr score ≥3 at baseline; calculated by guideline developers from patient-level data supplied by investigators (NB HY score ≥3 was a prespecified subgroup in the trial protocol and a randomisation stratification variable)

E.6.4 Deep brain stimulation compared with best medical treatment for earlier Parkinson's disease

E.6.4.1 Adverse events

		Quality assessment No. of events / no. of patients or patient-year		atient-years					
No. of studies	Design	Risk of bias	Incons- istency	Indirectness	Imprecision	DBS	Control	Effect (95% CI)	Quality
Serious	adverse	events (prob	ability of e	xperiencing ≥1)); 24 months				
1 ²	RCT	No serious	N/A	Not serious	Serious ³	68/124	56/127	RR = 1.24 (0.97 to 1.60)	MODERATE
Serious	adverse	events (rate j	per patient	-year); 24 mont	ths				
1 ²	RCT	No serious	N/A	Not serious	Serious ⁴	123 per 246pt-yrs ⁵	128 per 249pt-yrs ⁵	IRR = 0.97 (0.76 to 1.25)	MODERATE
Falls (pr	obability	of experienc	ing ≥1); 24	months					
1 ²	RCT	No serious	N/A	Not serious	Serious ⁴	8/124	5/127	RR = 1.64 (0.55 to 4.87)	MODERATE
Falls (ra	te per pat	tient-year); 2	4 months						
1 ²	RCT	No serious	N/A	Not serious	Serious ⁴	11 per 246pt-yrs ⁵	5 per 249pt-yrs ⁵	IRR = 2.23 (0.77 to 6.41)	MODERATE
1 2 3	Schüpbach Schüpbach at a 95% co	2007 2013 onfidence level,	data are coi	nsistent with appre	eciable harm and i	no meaningful effect			

at a 95% confidence level, data are consistent with appreciable benefit, appreciable harm and no meaningful effect assuming dropouts withdrew at 1 year (i.e. halfway through 2-year follow-up) 4

5

E.6.4.2 Symptom severity

Quality ass	essment					Number	of patients	Effect	
Number of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	DBS	Control	Mean difference (MD) (95% CI)	Quality
Hoehn and	Yahr scor	e (off medicatio	on) (lower is bette	er); 3–12 months					
1 ⁴	RCT	No serious	N/A	No serious	No serious	85	95	-0.32 (-0.56 to -0.09)	HIGH
Daily 'on' ti	me withou	It troublesome	dysinkesias (higl	her is better); 24	months				
1 ²	RCT	No serious	N/A	No serious	No serious	105	110	1.90 (0.51 to 3.29)	HIGH
Daily 'off' ti	me (lower	is better); 12-2	24 months						
2 ^{2,3}	RCT	No serious	No serious	No serious	No serious	209	212	-1.70 (-2.35 to -1.06)	HIGH
UPDRS I (lo	wer is bet	tter); 12–24 mo	nths						
3 ^{2,4,5}	RCT	No serious	No serious	No serious	Serious ⁷	233	225	-0.01 (-0.34 to 0.32)	MODERATE
UPDRS II or	n (lower is	better); 12–24	months						
4 ^{1,2,3,5}	RCT	No serious	No serious	No serious	Serious ⁷	246	244	0.48 (-0.40 to 1.37)	MODERATE
UPDRS III o	n (lower is	s better); 12–24	1 months						
4 ^{1,2,3,5}	RCT	No serious	No serious	No serious	No serious	243	241	-3.21 (-4.49 to -1.93)	HIGH
UPDRS IV (lower is b	etter); 12–24 m	onths						
4 ^{1,2,4,5}	RCT	No serious	Serious ⁶	No serious	No serious	214	212	-4.68 (-6.75 to -2.61)	MODERATE
1 Sahi	unhach 200	7							

² Schupbach 2007 ² Schüpbach 2013

² Schüpbach 2013

³ PDSURG (subgroup with baseline HY<3); multiply imputed data; ANCOVA model with baseline score, Hoehn and Yahr status (<3 -v- ≥3) and treatment allocation as covariates of final score; calculated by guideline developers from patient-level data supplied by investigators

⁴ PDSURG (subgroup with baseline HY<3); observed cases; ANCOVA model with baseline score, Hoehn and Yahr status (<3 -v- ≥3) and treatment allocation as covariates of final score; calculated by guideline developers from patient-level data supplied by investigators

⁵ Charles 2014

⁶ *l*² greater than 40% with no obvious explanation for heterogeneity

⁷ at a 95% confidence level, data are consistent with appreciable benefit and no effect

				Mean Difference	Mean Difference	
Study or Subgroup M	ean Difference	SE	Weight	IV, Fixed, 95% CI	IV, Fixed, 95% CI	
3.5.1 12 months						
PDSURG (baseline HY less than 3) ANCOVA; MI*	-1.676	0.372	78.3%	-1.68 [-2.41, -0.95]		
PDSURG (baseline HY less than 3) OC†	-2.057	0.496326	0.0%	-2.06 [-3.03, -1.08]		
PDSURG (meeting EarlyStim inclusion criteria); OC§	-2.902	0.75342	0.0%	-2.90 [-4.38, -1.43]	•	
Subtotal (95% CI)			78.3%	-1.68 [-2.41, -0.95]	\bullet	
Heterogeneity: Not applicable						
Test for overall effect: $Z = 4.51 (P < 0.00001)$						
3.5.2 18 months						
Subtotal (95% CI)				Not estimable		
Heterogeneity: Not applicable						
Test for overall effect: Not applicable						
3.5.3 24 months						
Schuepbach 2013	-1.8	0.707103	21.7%	-1.80 [-3.19, -0.41]		
Subtotal (95% CI)			21.7%	-1.80 [-3.19, -0.41]		
Heterogeneity: Not applicable						
Test for overall effect: $Z = 2.55$ (P = 0.01)						
Total (95% CI)			100.0%	-1.70 [-2.35, -1.06]	◆	
Heterogeneity: Chi ² = 0.02, df = 1 (P = 0.88); l ² = 0%				-		
Test for overall effect: Z = 5.17 (P < 0.00001)					-4 -2 U 2 4	
Test for subgroup differences: $Chi^2 = 0.02$, df = 1 (P = 0.88	3), I ² = 0%					
* multiply imputed data; ANCOVA model with baseline score † observed cases; unadjusted mean difference; not included	e, Hoehn and Ya I in meta-analysi	hr status (< is – shown f	3 -v- ≥3) a or compa	and treatment allocatio rison purposes only	on as covariates of final score	

§ participants meeting key eligibility criteria for EarlyStim (age 18–60; disease duration ≥4 years; Hoehn and Yahr <3; improvement of 50% or more with dopaminergic medication on UPDRS-III); observed cases; unadjusted mean difference; not included in meta-analysis – shown for comparison purposes only</p>

DBS -v- medication alone: mean daily 'off' time – forest plot

				Mean Difference		Mean Differe	ence	
Study or Subgroup	Mean Difference	SE	Weight	IV, Fixed, 95% C	I	IV, Fixed, 95	% CI	
3.2.1 12 months								
PDSURG (baseline HY less than 3) ANCOVA; OC*	-0.20413	0.215309	59.7%	-0.20 [-0.63, 0.22]				
PDSURG (baseline HY less than 3) OC†	-0.327	0.227256	0.0%	-0.33 [-0.77, 0.12]				
PDSURG (meeting EarlyStim inclusion criteria); OC§ Subtotal (95% CI)	-0.425	0.305964	0.0% 59.7%	-0.42 [-1.02, 0.17] -0.20 [-0.63, 0.22]				
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.95 (P = 0.34)								
3.2.2 18 months								
Subtotal (95% CI)				Not estimable				
Heterogeneity: Not applicable								
Test for overall effect: Not applicable								
3.2.3 24 months								
Charles 2014	0.1	0.756009	4.8%	0.10 [-1.38, 1.58]			<u> </u>	
Schuepbach 2013	0.3	0.279478	35.4%	0.30 [-0.25, 0.85]				
Subtotal (95% CI)			40.3%	0.28 [-0.24, 0.79]				
Heterogeneity: Chi ² = 0.06, df = 1 (P = 0.80); l ² = 0%								
Test for overall effect: Z = 1.05 (P = 0.29)								
Total (95% CI)			100.0%	-0.01 [-0.34, 0.32]		•		
Heterogeneity: Chi ² = 2.06, df = 2 (P = 0.36); l ² = 3%					l	+		—
Test for overall effect: $Z = 0.06$ (P = 0.95)					-2	-1 0	1	2
Test for subgroup differences: $Chi^2 = 2.00$, df = 1 (P = 0	.16), I² = 50.1%				F	avours DBS Fav	ours medication	
* observed cases; ANCOVA model with baseline score, F † observed cases; unadjusted mean difference; not includ § participants meeting key eligibility criteria for EarlyStim	loehn and Yahr sta ded in meta-analysi (age 18–60: diseas	tus (<3 -v- ≧ s – shown f e duration ≧	≥3) and tre or compa ≥4 years; I	eatment allocation a rison purposes only Hoehn and Yahr <3:	s covariates c improvement	f final score t of 50% or more v	vith dopaminergic	:

medication on UPDRS-III); observed cases; unadjusted mean difference; not included in meta-analysis – shown for comparison purposes only

DBS -v- medication alone: UPDRS I – forest plot

				Mean Difference		Меа	n Difference		
Study or Subgroup	lean Difference	SE	Weight	IV, Fixed, 95% C		IV, F	ixed, 95% C	I	
3.2.1 12 months									
PDSURG (baseline HY less than 3) ANCOVA; MI*	0.595	0.779	33.6%	0.59 [-0.93, 2.12]					
PDSURG (baseline HY less than 3) OC†	0.81	0.802492	0.0%	0.81 [-0.76, 2.38]					
PDSURG (meeting EarlyStim inclusion criteria); OC§	1.04	1.110806	0.0%	1.04 [-1.14, 3.22]					
Subtotal (95% CI)			33.6%	0.59 [-0.93, 2.12]					
Heterogeneity: Not applicable									
Test for overall effect: Z = 0.76 (P = 0.44)									
3.2.2.18 months									
Schuenhach 2007	-0.2	1 236032	13 3%	-0.20[-2.62.2.22]					
Subtotal (95% CI)	-0.2	1.200902	13.3%	-0.20 [-2.62, 2.22]					
Heterogeneity: Not applicable									
Test for overall effect: $7 = 0.16$ (P = 0.87)									
3.2.3 24 months									
Charles 2014	1.8	2.489612	3.3%	1.80 [-3.08, 6.68]					
Schuepbach 2013	0.5	0.640315	49.8%	0.50 [-0.75, 1.75]					
Subtotal (95% CI)			53.0%	0.58 [-0.63, 1.80]					
Heterogeneity: Chi² = 0.26, df = 1 (P = 0.61); l² = 0%									
Test for overall effect: Z = 0.94 (P = 0.35)									
Total (95% CI)			100.0%	0.48 [-0.40, 1.37]			•		
Heterogeneity: Chi ² = 0.61, df = 3 (P = 0.90); l ² = 0%					 				———————————————————————————————————————
Test for overall effect: $Z = 1.07$ (P = 0.29)					-10	-5	0	5	10
Test for subgroup differences: $Chi^2 = 0.35$, df = 2 (P = 0.84)	4), l² = 0%					⊢avours L	BS Favours	medication	1
3.2.3 24 months Charles 2014 Schuepbach 2013 Subtotal (95% CI) Heterogeneity: Chi ² = 0.26, df = 1 (P = 0.61); l ² = 0% Test for overall effect: $Z = 0.94$ (P = 0.35) Total (95% CI) Heterogeneity: Chi ² = 0.61, df = 3 (P = 0.90); l ² = 0% Test for overall effect: $Z = 1.07$ (P = 0.29) Test for subgroup differences: Chi ² = 0.35, df = 2 (P = 0.8)	1.8 0.5 4), I ² = 0%	2.489612 0.640315	3.3% 49.8% 53.0%	1.80 [-3.08, 6.68] 0.50 [-0.75, 1.75] 0.58 [-0.63, 1.80] 0.48 [-0.40, 1.37]	-10	-5 Favours D	0 DBS Favours		

* multiply imputed data; ANCOVA model with baseline score, Hoehn and Yahr status (<3 -v- \geq 3) and treatment allocation as covariates of final score † observed cases; unadjusted mean difference; not included in meta-analysis – shown for comparison purposes only

§ participants meeting key eligibility criteria for EarlyStim (age 18–60; disease duration \geq 4 years; Hoehn and Yahr <3; improvement of 50% or more with dopaminergic medication on UPDRS-III); observed cases; unadjusted mean difference; not included in meta-analysis – shown for comparison purposes only

DBS -v- medication alone: UPDRS II (on) - forest plot

				Mean Difference		Me	an Difference	9	
Study or Subgroup	Mean Difference	SE	Weight	IV, Fixed, 95% C		IV,	Fixed, 95% 0		
3.3.1 12 months									
PDSURG (baseline HY less than 3) ANCOVA; MI*	-1.592	1.199	29.8%	-1.59 [-3.94, 0.76]			•		
PDSURG (baseline HY less than 3) OC†	-1.878	1.2187	0.0%	-1.88 [-4.27, 0.51]					
PDSURG (meeting EarlyStim inclusion criteria); OC§ Subtotal (95% CI)	-1.842	1.744699	0.0% 29.8%	-1.84 [-5.26, 1.58] -1.59 [-3.94, 0.76]					
Heterogeneity: Not applicable									
Test for overall effect: Z = 1.33 (P = 0.18)									
3.3.2 18 months									
Schuepbach 2007	-2.25	1.5612	17.6%	-2.25 [-5.31, 0.81]					
Subtotal (95% CI)			17.6%	-2.25 [-5.31, 0.81]					
Heterogeneity: Not applicable									
Test for overall effect: Z = 1.44 (P = 0.15)									
3.3.3 24 months									
Charles 2014	-3.3	4.474924	2.1%	-3.30 [-12.07, 5.47]	◀				
Schuepbach 2013	-4.5	0.921954	50.4%	-4.50 [-6.31, -2.69]					
Subtotal (95% CI)			52.6%	-4.45 [-6.22, -2.68]					
Heterogeneity: $Chi^2 = 0.07$, $df = 1$ (P = 0.79); $I^2 = 0\%$									
Test for overall effect: $Z = 4.93$ (P < 0.00001)									
Total (95% CI)			100.0%	-3.21 [-4.49, -1.93]		•	•		
Heterogeneity: $Chi^2 = 4.16$, df = 3 (P = 0.24); $I^2 = 28\%$					10				
Test for overall effect: Z = 4.90 (P < 0.00001)					-10	-5 Favours	U DBS Eavour	o s medicatio	י מנ
Test for subgroup differences: $Chi^2 = 4.09$, df = 2 (P = 0).13), I² = 51.1%					1 470413		e mealoalle	***

† observed cases; unadjusted mean difference; not included in meta-analysis – shown for comparison purposes only § participants meeting key eligibility criteria for EarlyStim (age 18–60; disease duration ≥4 years; Hoehn and Yahr <3; improvement of 50% or more with dopaminergic medication on UPDRS-III); observed cases; unadjusted mean difference; not included in meta-analysis – shown for comparison purposes only

DBS -v- medication alone: UPDRS III (on) - forest plot

				Mean Difference		Mean D	ifference	
Study or Subgroup	Mean Difference	SE	Weight	IV, Random, 95% CI		IV, Rando	om, 95% Cl	
3.4.1 12 months								
PDSURG (baseline HY less than 3) ANCOVA; OC*	-3.77109	0.47251	27.6%	-3.77 [-4.70, -2.84]				
PDSURG (baseline HY less than 3) OC†	-3.773	0.52987	0.0%	-3.77 [-4.81, -2.73]				
PDSURG (meeting EarlyStim inclusion criteria); OC§ Subtotal (95% CI)	-4.29	0.748125	0.0% 27.6%	-4.29 [-5.76, -2.82] -3.77 [-4.70, -2.84]		•		
Heterogeneity: Not applicable								
Test for overall effect: Z = 7.98 (P < 0.00001)								
3.4.2 18 months								
Schuepbach 2007	-9.7	1.118034	22.0%	-9.70 [-11.89, -7.51]		_		
Subtotal (95% CI)			22.0%	-9.70 [-11.89, -7.51]		•		
Heterogeneity: Not applicable								
Test for overall effect: Z = 8.68 (P < 0.00001)								
3.4.3 24 months								
Charles 2014	-1.6	1.06444	22.5%	-1.60 [-3.69, 0.49]			+	
Schuepbach 2013	-4.1	0.419094	27.9%	-4.10 [-4.92, -3.28]		-		
Subtotal (95% CI)			50.4%	-3.04 [-5.46, -0.62]				
Heterogeneity: Tau ² = 2.47; Chi ² = 4.78, df = 1 (P = 0.03	; l² = 79%							
Test for overall effect: Z = 2.46 (P = 0.01)								
Total (95% CI)			100.0%	-4.68 [-6.75, -2.61]		\bullet		
Heterogeneity: Tau ² = 3.84; Chi ² = 31.03, df = 3 (P < 0.0	0001); l² = 90%							+
Test for overall effect: Z = 4.42 (P < 0.00001)					-10	-o Favours DRS	v o Favours medicatio	1U m
Test for subgroup differences: $Chi^2 = 25.41$, df = 2 (P < C	.00001), I ² = 92.10	%						

† observed cases; unadjusted mean difference; not included in meta-analysis – shown for comparison purposes only § participants meeting key eligibility criteria for EarlyStim (age 18–60; disease duration ≥4 years; Hoehn and Yahr <3; improvement of 50% or more with dopaminergic medication on UPDRS-III); observed cases; unadjusted mean difference; not included in meta-analysis – shown for comparison purposes only

DBS -v- medication alone: UPDRS IV – forest plot

E.6.4.3 Neuropsychological outcomes

Quality	y assessme	nt				Numbe	r of patients		
Number of stud	er dies Desig	In Risk of bias	Inconsistency	Indirectness	Imprecision	DBS	Control	Mean difference (MD) (95% CI)	Quality
Cognit	tive functio	n (MDRS) (higher	is better); 18–24	months					
2 ^{1,2}	RCT	No serious	Not serious	Not serious	Serious ³	134	137	0.61 (-0.47 to 1.68)	MODERATE
Depres	ssion (Mon	gomery–Åsberg	depression scale) (lower is bette	er); 18–24 mon	ths			
2 ^{1,2}	RCT	No serious	Not serious	Not serious	Not serious	133	137	-2.66 (-4.11 to -1.20)	HIGH
1 2	Schüpbach	2007							

³ at a 95% confidence level, data are consistent with appreciable benefit, appreciable harm and no meaningful effect

		DBS		N	ledication			Mean Difference		Mea	an Differ	ence		
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Fixed, 95% CI		IV,	Fixed, 9	5% CI		
3.10.1 18 months														
Schuepbach 2007	-1	5.348889	10	-0.25	4.253658	10	6.4%	-0.75 [-4.99, 3.49]			•			
Subtotal (95% CI)			10			10	6.4%	-0.75 [-4.99, 3.49]						
Heterogeneity: Not ap	plicable													
Test for overall effect:	Z = 0.35	5 (P = 0.73)												
3.10.2 24 months														
Schuepbach 2013	1.3	4.4542	124	0.6	4.5078	127	93.6%	0.70 [-0.41, 1.81]			-+-			
Subtotal (95% CI)			124			127	93.6%	0.70 [-0.41, 1.81]						
Heterogeneity: Not ap	plicable													
Test for overall effect:	Z = 1.24	(P = 0.22)												
Total (95% CI)			134			137	100.0%	0.61 [-0.47, 1.68]						
Heterogeneity: Chi ² =	0.42, df	= 1 (P = 0.5	52); l² =	0%							<u> </u>	<u> </u>		
Test for overall effect:	Z = 1.11	(P = 0.27)							-4	-2 -2		2	4 diaction	
		$Chi^2 = 0.42$	df - 1	$(\mathbf{P} = 0)$	$= 2) ^2 = 00/$				F	-avours L	льз га	vours me	edication	

DBS -v- medication alone: cognitive function (MDRS) – forest plot

		DBS		N	ledication			Mean Difference		Me	an Differen	се	
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Fixed, 95% C	I	IV,	Fixed, 95%	CI	
3.11.1 18 months													
Schuepbach 2007	-2.75	3.192602	10	0.75	3.686435	10	23.2%	-3.50 [-6.52, -0.48]					
Subtotal (95% CI)			10			10	23.2%	-3.50 [-6.52, -0.48]					
Heterogeneity: Not app	plicable												
Test for overall effect:	Z = 2.27	(P = 0.02)											
3.11.2 24 months													
Schuepbach 2013	-1.1	6.6543	123	1.3	6.7617	127	76.8%	-2.40 [-4.06, -0.74]		—			
Subtotal (95% CI)			123			127	76.8%	-2.40 [-4.06, -0.74]					
Heterogeneity: Not app	plicable												
Test for overall effect:	Z = 2.83	(P = 0.005)										
Total (95% CI)			133			137	100.0%	-2.66 [-4.11, -1.20]		-			
Heterogeneity: Chi ² =	0.39, df :	= 1 (P = 0.5	3); I² =	0%					+			<u> </u>	
Test for overall effect:	Z = 3.57	(P = 0.000	4)						-10	-5		5 ura madiaat	10 ion
Test for subgroup diffe	erences:	Chi ² = 0.39	, df = 1	(P = 0.8)	53), I² = 0%					Favours	DDS Favo	urs medical	

DBS -v- medication alone: depression (MADRS) – forest plot

E.6.4.4 Health related quality of life – patient

Quality as	sessment					Number	r of patients	Effect	
Number of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	DBS	Control	Mean difference (MD) (95% Cl)	Quality
EQ-5D (hig	gher is bett	ter); 12 months							
1 ³	RCT	No serious	NA	No serious	Very serious ⁴	104	99	0.00 (-0.08 to 0.07)	LOW
PDQ-39 (lo	ower is bet	ter); 12–24 mor	nths						
4 ^{1,2,3,5}	RCT	No serious	No serious	No serious	No serious	306	288	-5.96 (-8.27 to -3.65)	HIGH
¹ Sch ² Sch ³ PD	hüpbach 200 hüpbach 201 SURG (subg	7 3 proup with baseling re: calculated by	e HY<3); multiply imp	outed data; ANCOV	A model with base	eline score,	Hoehn and Yal	hr status (<3 -v- ≥3) and treat	ment allocation

as covariates of final score; calculated by guideline developers from patient-level data supplied by investigators 4 at a 95% confidence level, data are consistent with appreciable benefit, appreciable harm and no meaningful effect

Quality ass	essment					Number	of patients	Effect	
Number of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	DBS	Control	Mean difference (MD) (95% Cl)	Quality
5 Cha	rles 2014								

				Mean Difference		Меа	an Differenc	e	
Study or Subgroup	Mean Difference	SE	Weight	IV, Fixed, 95% C	I	IV,	Fixed, 95%	CI	
3.6.1 12 months									
PDSURG (baseline HY less than 3) ANCOVA; MI*	-3.35	1.999	34.8%	-3.35 [-7.27, 0.57]		—	■		
PDSURG (baseline HY less than 3) OC†	-3.49	1.659078	0.0%	-3.49 [-6.74, -0.24]					
PDSURG (meeting EarlyStim inclusion criteria); OC§	-4.021	2.685906	0.0%	-4.02 [-9.29, 1.24]					
Subtotal (95% CI)			34.8%	-3.35 [-7.27, 0.57]					
Heterogeneity: Not applicable									
Test for overall effect: Z = 1.68 (P = 0.09)									
3.6.2 18 months									
Schuenbach 2007	-8 1	5 210216	5 1%	-8 10 [-18 31 2 11]					
Subtotal (95% CI)	0.1	0.210210	5.1%	-8.10 [-18.31, 2.11]					
Heterogeneity: Not applicable									
Test for overall effect: $Z = 1.55$ (P = 0.12)									
3.6.3 24 months									
Charles 2014	-2.4	4.27614	7.6%	-2.40 [-10.78, 5.98]			•	-	
Schuepbach 2013	-8	1.627882	52.5%	-8.00 [-11.19, -4.81]					
Subtotal (95% CI)			60.1%	-7.29 [-10.27, -4.31]		\bullet			
Heterogeneity: Chi ² = 1.50, df = 1 (P = 0.22); l ² = 33%									
Test for overall effect: Z = 4.79 (P < 0.00001)									
Total (95% CI)			100.0%	-5.96 [-8.27, -3.65]		•			
Heterogeneity: Chi ² = 4.14, df = 3 (P = 0.25); l ² = 27%					I				
Test for overall effect: $Z = 5.06 (P < 0.00001)$					-20	-10	0	10	20
Test for subgroup differences: $Chi^2 = 2.64$, $df = 2$ (P = 0).27), l² = 24.2%					Favours L	JBS Favou	is medication	11

* multiply imputed data; ANCOVA model with baseline score, Hoehn and Yahr status (<3 -v- \geq 3) and treatment allocation as covariates of final score † observed cases; unadjusted mean difference; not included in meta-analysis – shown for comparison purposes only δ participants meeting key eligibility criteria for FadyStim (age 18–60; disease duration >4 years; Hoehn and Yahr <3; improvement of 50% or more with d

§ participants meeting key eligibility criteria for EarlyStim (age 18–60; disease duration ≥4 years; Hoehn and Yahr <3; improvement of 50% or more with dopaminergic medication on UPDRS-III); observed cases; unadjusted mean difference; not included in meta-analysis – shown for comparison purposes only

DBS -v- medication alone: PDQ-39 – forest plot

E.6.4.5 Medication load

Quality	asses	ssment					Number	of patients	Effect	
Numbe of stud	er lies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	DBS	Control	Mean difference (MD) (95% Cl)	Quality
Daily d	osage	of anti-Pa	arkinson's med	lication (levodopa	mg equivalent)	(lower is better); 24 mon	ths		
3 ^{1,2,3}		RCT	No serious	Serious ⁴	No serious	No serious	149	151	-469 (-765 to -173)	MODERATE
1	Schüp	bach 2007								
2	Schüp	bach 2013								
3	Charle	es 2014								
4	l ² grea	ater than 40	% with no obvious	explanation for hete	rogeneity					

		DBS		N	ledication			Mean Difference		Mea	n Differen	се	
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Random, 95% C		IV, Ra	andom, 95	% CI	
3.9.1 12 months													
Subtotal (95% CI)			0			0		Not estimable					
Heterogeneity: Not app	olicable												
Test for overall effect:	Not appli	cable											
3.9.2 18 months													
Schuepbach 2007	-520	364.6917	10	130	232.9163	10	29.9%	-650.00 [-918.20, -381.80]					
Subtotal (95% CI)			10			10	29.9%	-650.00 [-918.20, -381.80]					
Heterogeneity: Not app	olicable												
Test for overall effect:	Z = 4.75	(P < 0.0000)1)										
3.9.3 24 months													
Charles 2014	97.7	344.7208	15	214.5	342.234	14	30.9%	-116.80 [-366.94, 133.34]					
Schuepbach 2013	-363.3	216.0293	124	245.8	211.8652	127	39.2%	-609.10 [-662.05, -556.15]					
Subtotal (95% CI)			139			141	70.1%	-378.75 [-860.20, 102.70]					
Heterogeneity: Tau ² =	112670.6	56; Chi² = 1	4.24, d [.]	f = 1 (P	= 0.0002); I	² = 93%	, D						
Test for overall effect:	Z = 1.54	(P = 0.12)											
Total (95% CI)			149			151	100.0%	-469.17 [-765.21, -173.13]					
Heterogeneity: Tau ² =	57519.17	7; Chi² = 14	.44, df	= 2 (P =	0.0007); l ²	= 86%							400
Test for overall effect:	Z = 3.11	(P = 0.002)							-1000	-500 Eavoura F		500	100 tion
Test for subaroup diffe	rences: (Chi ² = 0.93.	df = 1	(P = 0.3	3), I² = 0%						DS Favo	urs medical	lion

DBS -v- medication alone: medication load (levodopa equivalent mg/day)

E.7 Managing and monitoring impulse control disorder as an adverse effect of dopaminergic treatment

E.7.1 Predictors for the development of impulse control disorders

Quality assessment						Number of patients		Effect			
Number of studies	Design	Risk of bias	Inconsiste ncy	Indirectnes s	Imprecisio n	With ICD	No ICD	OR: 95%CI	Quality		
Male gender	Male gender										
Joutsa 2012 Antonini 2016	Cohort	Serious ¹	Serious ²	Not serious ³	Not serious	93	963	6.10 (2.16 to 17.18) 1.14 (0.68, 1.92)	LOW		
Comorbid an	ixiety or de	pression									
Pontone 2006	Cohort	Very serious ⁴	N/A ⁶	Not serious ³	Serious ⁵	9	100	2.54 (0.6 to 10.15)	VERY LOW		
DA use											
Pontone 2006 Voon 2007	Cohort	Very serious ⁴	Not serious	Not serious ³	Not serious	30	386	10.46 (3.13 to 34.91)	LOW		
Pramipexole use											
Imamura 2008 Pontone 2006 Sharma 2015	Cohort	Very serious ⁴	Not serious	Not serious ³	Not serious	20	137	3.26 (1.99 to 5.35)	LOW		

Predictive factors for the development of ICD - unadjusted odds ratios (OR)

¹ Unadjusted odds ratio

² N/A; not applicable as only 1 study contributed to this analysis

³ No serious indirectness; population is as described in review protocol

⁴ Serious risk of bias, as assessed by NICE or CASP quality assessment checklist and unadjusted odds ratios

⁵ Non-significant results

⁶ Serious inconsistency; confidence intervals around point estimates do not overlap

Quality assessment						Number of patients		Effect		
Number of studies	Design	Risk of bias	Inconsiste ncy	Indirectnes s	Imprecisio n	With ICD	No ICD	OR: 95%CI	Quality	
Amantadine	use									
Weintraub 2010b Sharma 2015	Cohort	Serious ¹	Not serious	Not serious	Not serious	728	2357	1.68 (1.36 to 2.08)	MODERATE	
Levodopa us	se									
Imamura 2008 Antonini 2016	Cohort	Serious ¹	Not serious	Serious ³	Serious⁵	82	752	0.27 (0.05 to 1.29) 2.35 (0.83 to 6.61)	VERY LOW	
Combination	ı levodopa	and pramipes	kole therapy							
lmamura 2008	Cohort	Serious ¹	N/A ²	Serious ³	Serious ⁵	11	37	1.96 (0.3 to 8.79)	VERY LOW	
Entacapone	use									
Sharma 2015	Cohort	Serious ¹	N/A ²	Not serious	Serious ⁵	74	255	1.47 (0.75 to 2.9)	LOW	
Rasagaline u	use									
Sharma 2015	Cohort	Serious ¹	N/A ²	Not serious	Serious ⁵	74	255	0.98 (0.5 to 1.9)	LOW	
Marriage status (unmarried)										
Sharma 2015	Cohort	Serious ¹	N/A ²	Not serious	Not serious	74	255	9.6 (2.9 to 31.3)	MODERATE	
Alcohol intake (high alcohol consumption)										
Sharma 2015	Cohort	Serious ¹	N/A ²	Not serious	Not serious	74	255	4.0 (2.0 to 8.05)	MODERATE	
Smoker status (smoker)										
Imamura 2008	Cohort	Serious ¹	N/A ²	Serious ³	Not serious	11	37	7.5 (3.5 to 16.15)	LOW	

 ¹ Unadjusted odds ratio
 ² N/A; not applicable as only 1 study contributed to this analysis
 ³ Serious indirectness; population was comprised of only those with pathological gambling

Quality assessment						Number of patients		Effect		
Number of studies	Design	Risk of bias	Inconsiste ncy	Indirectnes s	Imprecisio n	With ICD	No ICD	OR: 95%CI	Quality	
Family history of alcohol or gambling abuse										
Voon 2007	Cohort	Serious ¹	N/A	Not serious	Not serious	21	286	5.66 (1.78 to 18.03)	MODERATE	
Rotigotine dose (12-16mg/day versus 2-10mg/day)										
Antonini 2016	Cohort	Serious ¹	N/A	Not serious	Serious ⁵	71	715	0.66 (0.40 to 1.08)	LOW	

Quality assessment						Number of patients		Effect		
Number of studies	Design	Risk of bias	Inconsiste ncy	Indirectnes s	Imprecisio n	With ICD	No ICD	Adjusted OR (95%CI)	Quality	
Younger age	e at onset o	of PD								
4 studies: Auyeung 2011 Gliadi 2007 Wentraub 2006 Sharma 2015	Cohort	Serious ¹	Serious ²	Not serious	Not serious	844	2976	OR1: 4.1 (1.1 to 15.9) OR2: 0.99 (0.99 to 1.00) OR3: 2.40 (1.91 to 3.02) OR4: 0.96 (0.93 to 0.99)	LOW	
Comorbid ar	nxiety or de	epression								
Auyeung 2011	Cohort	Serious ³	N/A	Not serious	Not serious	15	198	10.0 (2.0 to 50.8)	MODERATE	
Joutsa 2012	Cohort	Not serious	N/A	Not serious	Not serious	22	248	1.095 (1.001 to 1.195)	HIGH	
Gender male)									
2 studies: Gliadi 2007 Weintraub 2006	Cohort	Serious ¹	N/A	Not serious	Serious ⁵	782	2689	OR1: 1.10 (1.00 to 1.22) OR2: 4.34 (0.54 to 34.4871)	LOW	
DA use										
2 studies: Weintraub 2006 Weintraub 2010a	Cohort	Not serious	Not serious	Not serious	Not serious	749	2608	OR1: 16.7 (2.61 to 100) OR2: 2.64 (2.01 to 3.46)	HIGH	

Predictive factors for the development of ICD - Adjusted odds ratios (OR)

 ¹ Serious risk of bias as assessed by CASP cohort study checklist. Due to the very tight confidence intervals, this Gliadi et al study is heavily weighing the overall estimate
 ² Serious inconsistency; confidence intervals around point estimates do not overlap
 ³ Serious risk of bias: Study unclear as to how depression is retrospectively accounted for an in what subset of the study population
$^{^{\}rm 1}$ Serious risk of bias, as assessed by CASP cohort study quality checklist $^{\rm 2}$ Non-significant results

DA LEDD 60	-160 mg/d	l							
Lee 2010	Cohort	Not serious ¹	Not serious	Not serious ²	Not serious	118	1049	3.3 (1.3 to 9.1)	HIGH
DA LEDD >	• 150mg/da	ay							
Lee 2010 Sharma 2015	Cohort	Not serious ¹	Serious ³	Not serious ³	Not serious	118	1049	OR1 = 4.3 (1.6 to 11.9) OR2 = 4.52 (1.6 to 12.5)	MODERATE
DA LEDD 4	00 - 800m	g/day							
Lee 2010 Sharma 2015	Cohort	Not serious ¹	Serious ⁴	Not serious ³	Serious ⁸	118	1049	OR1 = 0.8 (0.4 to 1.6) OR2 = 1.38 (0.5 to 3.82)	LOW
DA LEDD >	>750mg/da	ay							
Lee 2010	Cohort	Not serious ¹	N/A ⁴	Not serious ³	Serious ⁸	118	1049	1.0 (0.5 to 2.1)	MODERATE
DA treatmen	t duration	< 2 years							
Gliadi 2007	Cohort	Serious ⁵	N/A ⁵	Serious ⁶	Serious ⁸	27	166	0.95 (0.84 to 1.08)	VERY LOW
DA treatmen	t duration	3 - 5 years							
Gliadi 2007	Cohort	Serious ⁶	N/A ⁵	Serious ⁷	Serious ⁸	27	166	1.04 (0.01 to 1.18)	VERY LOW
DA treatmen	nt duration	> 6 years							
Gliadi 2007	Cohort	Serious ⁶	N/A ⁵	Serious ⁷	Not serious	27	166	1.18 (1.00 to 1.39)	LOW
Amantadine	use								
2 studies: Weintraub 2006/2010a	Cohort	Not serious ¹	Not serious	Not serious ³	Not serious	749	2608	1.35 (1.07 to 1.70)	HIGH

¹ Low risk of bias, as assessed by CASP cohort study quality check list

 ² No serious indirectness; population was as described in review protocol
³ Serious inconsistency: Lee and Sharma define drug dosage differently, whereby Lee defined >160mg and 540-750mg; Sharma defines as 150-300mg, and >300mg
⁴ NA; not applicable as one only study contributed to this analysis
⁵ Serious risk of bias, as assessed by CASP cohort study quality check list
⁶ Serious indirectness; population was comprised of those with CGEC behaviours, not ICD diagnosis
⁸ Non-significant results

Levodopa use										
Weintraub 2010a	Cohort	Not serious ¹	N/A ²	Not serious ³	Not serious	728	2357	1.51 (1.09 to 2.09)	HIGH	
Prior history	of ICD sy	mptoms								
Weintraub 2006	Cohort	Not serious ¹	N/A ²	Not serious ³	Not serious	21	251	15.54 (2.83 to 76.16)	HIGH	
Family histo	Family history of alcohol abuse									
Weintraub 2010a	Cohort	Not serious ¹	N/A ²	Not serious ³	Not serious	728	2357	2.08 (1.33 to 3.25)	HIGH	

Incidence of ICD

Quality assess	ment					Number of patier		
Number of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	ICD	No ICD	Quality
ICD rate with s	hort- and long	J-acting DAs						
Rizos 2016	Survey based on medical records and clinical interviews	Not serious	N/A	Not serious	Serious ⁴	57	368	MODERATE
Incidence of ICD and association with dopamine replacement therapy								
Wang 2016	Interviews	Not serious	N/A	Not serious	Serious ⁴	9	208	MODERATE

 ¹ Low risk of bias, as assessed by CASP study quality checklist
² NA; not applicable as only one study contributed to the analysis
³ No serious indirectness; population was as described in review protocol
⁴ Serious imprecision: Low numbers of ICD vs no ICD

E.7.2 Managing dopaminergic treatment in people who have developed impulse control disorder

Adjustment of dopaminergic medication

Quality asse	ssment					Number of patients	Effect	
Number of studies	Design	Risk of bias	Inconsiste ncy	Indirectnes s	Imprecisio n	Patients with ICD (N=18)	n/N (%) resolution of symptoms	Quality
Discontinuat	tion of dopam	inergic therap	у					
Bastiaens 2013	Cohort	Not serious ¹	N/A ²	Not serious ³	Very serious⁴	n=10	10/10 (100%)	LOW
Reduction of	f dopaminergi	c therapy						
Bastiaens 2013	Cohort	Not serious ¹	N/A ²	Not serious ³	Very Serious⁴	n=5	3/5 (60%)	LOW
Continue sa	me dosage of	dopaminergio	therapy					
Bastiaens 2013	Cohort	Not serious ¹	N/A ²	Not serious ³	Very Serious ⁴	n=3	0/3 (0%)	LOW
Number of studies	Design	Risk of bias	Inconsiste ncy	Indirectnes s	Imprecisio n	Patients with ICD (N=18)	n/N with DAWS	Quality
Developmen	t of DAWS in	those who dis	continued do	paminergic th	erapy			
Bastiaens 2013	Cohort	Not serious ¹	N/A ²	Not serious ³	Very serious ⁴	10	4/10	LOW
Developmen	t of DAWS in	those who rec	luced dopami	nergic therapy	/			
Bastiaens 2013	Cohort	Not serious ¹	N/A ²	Not serious ³	Very Serious⁴	5	1/5	LOW
Developmen	t of DAWS in	those who co	ntinued same	dosage of do	paminergic the	erapy		
Bastiaens 2013	Cohort	Not serious ¹	N/A ²	Not serious ³	Very Serious ⁴	3	1/3	LOW

 ¹ Low risk of bias, as assessed by CASP cohort study quality checklist
² NA; not applicable, only 1 study contributed to this analysis
³ No serious indirectness, study population were as outlined in review protocol

⁴ Very serious imprecision; very small sample size to derive meaningful population prevalence estimates

Cognitive behavioural therapy (CBT) for ICD

Quality asse	ssment					Number of patients		Effect	
Number of studies	Design	Risk of bias	Inconsiste ncy	Indirectnes s	Imprecisio n	СВТ	Control	MD: 95%CI	Quality
Resolution of	of ICD sympto	ms							
Okai 2013	RCT	Not serious ¹	N/A ²	Not serious ³	Not serious	27	17	-4.17 (-5.8 to -2.5)	HIGH
Effect of CB	T on CGIC sco	ore							
Okai 2013	RCT	Not serious ¹	N/A ²	Not serious ³	Not serious	27	17	-0.8 (-5.6 to -0.3)	HIGH
Effect of CB	T on general h	ealth (GHQ)							
Okai 2013	RCT	Not serious ¹	N/A ²	Not serious ³	Not serious	27	17	-3.8 (-5.6 to -2.0)	HIGH
Effect of CB	T on mental h	ealth (NPI)							
Okai 2013	RCT	Not serious ¹	N/A ²	Not serious ³	Not serious	27	17	-4.7 (-9.1 to -0.3)	HIGH
Effect of CB	T on social ad	justment							
Okai 2013	RCT	Not serious ¹	N/A ²	Not serious ³	Not serious	27	17	-3.6 (-6 to -1.3)	HIGH
Effect of CB	T on depressi	on (BDI)							
Okai 2013	RCT	Not serious ¹	N/A ²	Not serious ³	Serious ⁴	27	17	-3.5 (-6.6 to 0.4)	MODERATE
Effect of CB	T on anxiety (I	BAI)							
Okai 2013	RCT	Not serious ¹	N/A ²	Not serious ³	Serious ⁴	27	17	-1.8 (-5.4 to 1.8)	MODERATE
Effect of CB	T on carers pe	erception of th	ne quality of th	eir relationshi	ip with their p	artner (GRIMS	marital state)	
Okai 2013	RCT	Not serious ¹	N/A ²	Not serious ³	Serious ⁴	27	17	-2.3 (-5.7 to 1.3)	MODERATE
Effect of CB	T on carers ge	eneral health ((GHQ)						
Okai 2013	RCT	Not serious ¹	N/A ²	Not serious ³	Serious ⁴	27	17	-1.5 (-3.2 to 0.1)	MODERATE

¹ Low risk of bias, as assessed by NICE RCT study quality checklist ² NA; not applicable, only 1 study contributed to this analysis

Quality assessment							oatients	Effect	
Number of studies	Design	Risk of bias	Inconsiste ncy	Indirectnes s	Imprecisio n	СВТ	Control	MD: 95%CI	Quality
¹ Low risk of bias, as assessed by NICE RCY study quality checklist									

²NA; not applicable, only 1 study contributed to this analysis

³ No serious indirectness, study population were as outlined in review protocol

⁴ Non-significant results

Naltrexone therapy

Quality asse	ssment					Number of patients		Effect			
Number of studies	Design	Risk of bias	Inconsiste ncy	Indirectnes s	Imprecisio n	Naltrexone	placebo	MD: 95%CI	Quality		
QUIP ICD sc	ore										
Papay 2014	RCT	Not serious ¹	N/A ²	Not serious ³	Not serious	24	26	7.37 (2.45 to 12.66)	HIGH		
Change in C	Change in CGIC score (change of 1 or 2 points from baseline)										
Papay 2014	RCT	Not serious ¹	N/A ²	Not serious ³	Serious ⁴	24	26	OR = 1.57 (0.47 to 5.23)	MODERATE		
UPDRS moto	or sore										
Papay 2014	RCT	Not serious ¹	N/A ²	Not serious ³	Serious ⁵	24	26	-3.70 (-9.24 to 1.84)	MODERATE		
Adverse eve	ents that lead t	o study disco	ntinuation								
Papay 2014	RCT	Not serious ¹	N/A ²	Not serious ³	Not serious	24	26	0	LOW		
¹ Low risk of bi ² N/A; not appl	Low risk of bias, as assessed by NICE RCT study quality checklist ? N/A; not applicable, only 1 study contributed to this analysis										

³ No serious indirectness, study population were as outlined in review protocol

⁴ Non-significant result ⁵ CI cross the MID between 3.25 (Horvath et al., 2015) and 5 points (Schrag et al., 2006)

Amantadine therapy

Quality asse	ssment					Number of patients		Effect	
Number of studies	Design	Risk of bias	Inconsiste ncy	Indirectnes s	Imprecisio n	Amantadin e	placebo	MD (95% CI)	Quality
Symptom ass	essment scale	(SAS)							
Thomas 2010	Cross-over RCT	Serious ¹	N/A ²	Serious ³	Not serious	12	5	-9.6 (-10.12 to-9.08)	LOW
Yale-Brown o	bsessive comp	oulsive scale (Y	-BOCS)						
Thomas 2010	Cross-over RCT	Serious ¹	N/A ²	Serious ³	Not serious	12	5	-9.17 (-11.1 to -10.3)	LOW
Resolution of	PG spending b	oehaviour							
Thomas 2010	Cross-over RCT	Serious ¹	N/A ²	Serious ³	Not serious	12	5	-16.40 (-18.73 to -14.27)	LOW
Adverse ever	nts								
Thomas 2010	Cross-over RCT	Serious ¹	N/A ²	Serious ³	Not serious	12	5	5 patients dropped out of the amantadine group	LOW

¹ Serious risk of bias, as assessed by NICE RCT quality checklist
² N/A; not applicable as only 1 study contributed to this analysis
³ Serious indirectness; population was composed of those with pathological gambling only

 ¹ Serious risk of bias, as assessed by NICE RCT quality checklist
² NA; not applicable as only 1 study contributed to this analysis
³ Serious indirectness; population was composed of those with pathological gambling only

E.8 Palliative Care

Patient support needs

Quality asse Number of p	essment atients							
Example	Studies	Design	Risk of bias	Inconsiste ncy	Indirectness	N	Score on support need survey; 0 (no need) to 5 (serious need)	Quality
Highest self	-rated support	needs of pati	ients with PD (
Information about PD	Kirstjanson (2006)	Survey	Serious ¹	N/A ²	Serious ³	174	3.5	LOW
Equipment for daily living	Kirstjanson (2006)	Survey	Serious ¹	N/A ²	Serious ³	174	2.62	LOW

Need for open discussion concerning treatment and care

Quality asse Number of p	ssment atients							
Example	Studies	Design	Risk of bias	Inconsiste ncy	Indirectness	N	Supporting statement	Quality
Open dialog	ue between p	atient and clin	ician					
Discussion of medication	Giles (2009)	Interview	Very serious ⁴	N/A ²	Serious⁵	2	"I'm the type of woman, I'm afraid to ask too many questions because sometimes I feel like they would say, like you're asking too many questions, just take the pills" (from Giles et al., 2009)	VERY LOW

¹ Serious risk of bias (CASP cohort quality check list): Methodology not clear, not clear whether all survey material was standardised or validated

² N/A; not applicable, single study

 ³ Serious indirectness - population was restricted to moderate disease; no advanced or newly diagnosed participants
⁴ Very serious risk of bias (CASP qualitative check list): study methodology unclear, interview open to researcher interpretation, role of interviewer in shaping response unclear

⁵ Serious indirectness; very small number of patients,

Advance care directives

Quality asse Number of p	ssment atients							
Example	Studies	Design	Risk of bias	Inconsiste ncy	Indirectness	N	Supporting statement	Quality
Advanced ca	are directives							
Input from healthcare team to inform planning	2: Giles (2009) Hasson (2010)	Interview	Very serious ¹	Not serious ²	Not serious ³	22	"To help the family or as a group decide what would be the best care situation for the person, and you know what to expect" (from Giles et al., 2009)	LOW

Advance care planning

Quality asse Number of p	essment patients							
Example	Studies	Design	Risk of bias	Inconsiste ncy	Indirectnes s	N	Percentage(%) of patients who completed action	Quality
Advanced p	lanning of leg	al will						
Complete will	Kwak (2014)	Survey	Serious ⁴	N/A ⁵	Not serious ³	64	93.7%	MODERATE
Share will with spouse	Kwak (2014)	Survey	Serious ⁴	N/A ⁵	Not serious ³	64	90.6%	MODERATE
Share will with physician	Kwak (2014)	Survey	Serious ⁴	N/A ⁵	Not serious ³	64	37.5%	MODERATE
Preferences	for communi	cation about a	dvance care	planning				
Advance care	Tuck (2015)	Survey	Serious ⁴	N/A ⁵	Not serious ³	267	68.5% (with any kind of advance care planning documents)	MODERATE

¹ Very serious risk of bias (CASP qualitative check list); Hasson (2010) study was retrospective and open to memory bias; methodology very open to researcher interpretation and unclear in Giles (2009)

⁵ N/A, not applicable; single study

 ² No serious inconsistency, both studies share similar message
³ No serious indirectness, all participants were carers of a person with PD and therefore matched protocol
⁴ Serious risk of bias (CASP cohort quality check list): Methodology not clear, not clear whether all survey/questionnaire material was standardised or validated

Quality asse Number of p	essment atients							
Example	Studies	Design	Risk of bias	Inconsiste ncy	Indirectnes s	N	Percentage(%) of patients who completed action	Quality
planning documents								
When should your doctor discuss advance care planning	Tuck (2015)	Survey	Serious ¹	N/A ²	Not serious	267	-	MODERATE
Who should ideally raise issues regarding advance care planning to discuss	Tuck (2015)	Survey	Serious ¹	N/A ²	Not serious	267	94.4% responded	MODERATE

Support needs

Quality asse Number of p	ssment atients							
Example	Studies	Design	Risk of bias	Inconsiste ncy	Indirectness	N	Score on support need survey; 0 (no need) to 5 (serious need)	Quality
Greatest sup	port needs ider	ntified by carers	s (mean score :					
Information: how to provide care	Kirstjanson (2006)	Survey	Serious ¹	N/A ²	Serious ³	141	3.31	LOW

 ¹ Serious risk of bias (CASP cohort quality check list): Methodology not clear, not clear whether all survey material was standardised or validated
² N/A; not applicable, single study
³ Serious indirectness - population was restricted to moderate disease; no advanced or newly diagnosed participants

Quality asse Number of p	essment atients							
Example	Studies	Design	Risk of bias	Inconsiste ncy	Indirectness	N	Score on support need survey; 0 (no need) to 5 (serious need)	Quality
Reliable support workers	Kirstjanson (2006)	Survey	Serious ¹	N/A ²	Serious ³	141	2.84	LOW
Financial assistance for care	Kirstjanson (2006)	Survey	Serious ¹	N/A ²	Serious ³	141	2.72	LOW
Flexible home support program access	Kirstjanson (2006)	Survey	Serious ¹	N/A ²	Serious ³	141	2.52	LOW

Multidisciplinary care

Quality asse Number of p	essment atients							
Example	Studies	Design	Risk of bias	Inconsiste ncy	Indirectness	N	Supporting statement	Quality
Multidisciplina	ary care need							
Need for coordinated care	2: Hasson (2010) Giles (2009)	Interview	Very serious ¹	Not serious ²	Not serious ³	22	"There seems to be a vague boundary between the responsibilities that one person has and the responsibilities another has. They just don't seem to work as a team or have any team effort as such. You are nearly taking pot luck with each one in tum" (Giles et al., 2009)	LOW

 ¹ Very serious risk of bias (CASP qualitative check list); Hasson study was retrospective and open to memory bias; methodology very open to researcher interpretation and unclear in Giles (09)
² No serious inconsistency, both studies share similar message
³ No serious indirectness, all participants were carers of a person with PD and therefore matched protocol

Decision making

Quality asses	sment							
Number of pa	tients		Diels of	Inconciete				
Example	Studies	Design	bias	ncy	Indirectness	N	care goal	Quality
End of life ca	re goals							
Several people discuss; 1 person decide on action	Kwak (2014)	Survey	Serious ¹	N/A ²	Not serious ³	64	53%	MODERATE
One person decide alone	Kwak (2014)	Survey	Serious ¹	N/A ²	Not serious ³	64	28%	MODERATE
Several people decide on action together	Kwak (2014)	Survey	Serious ¹	N/A ²	Not serious ³	64	14%	MODERATE
Carer to be involved in decision making	Kwak (2014)	Survey	Serious ¹	N/A ²	Not serious ³	64	92%	MODERATE
Other family members to be involved in decision making	Kwak (2014)	Survey	Serious ¹	N/A ²	Not serious ³	64	72%	MODERATE
Physician to be involved in decision making	Kwak (2014)	Survey	Serious ¹	N/A ²	Not serious ³	64	70%	MODERATE
Carer, family, and	Kwak (2014)	Survey	Serious ¹	N/A ²	Not serious ³	64	52%	MODERATE

¹ Serious risk of bias: Methodology not clear, not clear whether all survey material was standardised or validated ² N/A, single study

Quality assessment Number of patients								
Example	Studies	Design	Risk of bias	Inconsiste ncy	Indirectness	N	Percentage(%) of carers who elected care goal	Quality
physician to be involved in decision making								

Information needs

Quality asse Number of p	ssment atients								
Example	Studies	Design	Risk of bias	Inconsiste ncy	Indirectness	N	Supporting statement	Quality	
Information at diagnosis about Parkinson's disease									
understandi ng the disease	Giles (2009)	Interview	Very serious ¹	N/A ²	Serious ³	5	"you have to be prepared and understand it's just kind of a shocker and no one really explained to us what all of this meant" (from Giles et al., 2009)	VERY LOW	
Information to	help carers pi	repare to adva	ncement of dise	ease					
Preparation for end of life	Hasson (2010)	Interview	Serious ^₄	N/A ²	Not serious ⁵	15	<i>"I knew he was deteriorating but I didn't expect him to die so soon" (Hasson et al., 2010)"</i>	MODERATE	

 ¹ Very serious risk of bias (CASP qualitative check list): methodology unclear and open to researcher interpretation
² N/A, not applicable, single study
³ Serious indirectness, very small sample size, may be unrepresentative of general population
⁴ Serious bias (CASP qualitative check list), retrospective perspective may bias responses
⁵ No serious indirectness; carers of patient with PD as specified in protocol

Satisfaction with care

Quality asse Number of p	essment atients							
Example	Studies	Design	Risk of bias	Inconsiste ncy	Indirectness	N	Percentage (%) of carers who reported satisfaction (FAMCARE assessment)	Quality
Satisfaction	with care reco	eived						
Information giving	Kirstjanson (2006)	Survey	Serious ¹	N/A	Serious ²	141	69%	LOW
Physical care	Kirstjanson (2006)	Survey	Serious ¹	N/A	Serious ²	141	80%	LOW
Phycosocial care	Kirstjanson (2006)	Survey	Serious ¹	N/A	Serious ²	141	63%	LOW
Availability of care	Kirstjanson (2006)	Survey	Serious ¹	N/A	Serious ²	141	71%	LOW

Respite opportunities and availability of care

Quality assessment Number of patients								
Example	Studies	Design	Risk of bias	Inconsiste ncy	Indirectness	N	Supporting statement	Quality
Respite oppo	ortunities							
Access to respite	2: Hasson (2010) Giles (2009)	Interview	Very serious ³	Not serious ⁴	Not serious ⁵	22	"they (government homecare) still haven't called usso we're lucky that, you know, we finally made the decision to move on. Because I don't know what we would have done I don't think my mom would have lasted" (from Giles et al., 2009)	LOW

¹ Serious risk of bias: Methodology not clear, not clear whether all survey material was standardised or validated ² Serious indirectness - population was restricted to moderate disease; no advanced or newly diagnosed participants

³ Very serious risk of bias (CASP gualitative check list); Hasson (2010) study was retrospective and open to memory bias; methodology very open to researcher interpretation and unclear in Giles (2009)

⁴ No serious inconsistency, both studies share similar message

⁵ No serious indirectness, all participants were carers of a person with PD and therefore matched protocol

Quality assessment Number of patients								
Example	Studies	Design	Risk of bias	Inconsiste ncy	Indirectness	N	Supporting statement	Quality

Access to domiciliary palliative care services

Quality asse Number of p	ssment atients							
Example	Studies	Design	Risk of bias	Inconsiste ncy	Indirectness	N	Supporting statement	Quality
Access to do	miciliary palliati	ive care service	es					
Access to palliative care services	2: Hasson (2010) Giles (2009)	Interview	Very serious ⁴	Not serious ²	Not serious	22	"that (home care services) is something that you know somebody should tell those people". (from Giles et al., 2009)	LOW

Patient and carer quality of life

Quality asse Number of p	ssment atients							
Example	Studies	Design	Risk of bias	Inconsiste ncy	Indirectness	N	Mean score (SD) on self-rated QoL scale (0 = very poor, 10 = excellent\)	Quality
Patient quali	ty of life (QoL	.)						
Patient- rated QoL	Kirstjanson (2006)	Survey	Serious ¹	N/A ²	Serious ³	174	6.87 (2.29)	LOW
Satisfaction with QoL	Kirstjanson (2006)	Survey	Serious ¹	N/A ²	Serious ³	174	5.55 (2.68)	LOW

¹ Very serious risk of bias: Methodology not clear, not clear whether all survey material was standardised or validated

 $^{^{2}}$ N/A, not applicable, single study

 ³ Serious indirectness - population was restricted to moderate disease; no advanced or newly diagnosed participants
⁴ Very serious risk of bias (CASP qualitative check list); Hasson (2010) study was retrospective and open to memory bias; methodology very open to researcher interpretation and unclear in Giles (2009)

Quality assessment Number of patients								
Example	Studies	Design	Risk of bias	Inconsiste ncy	Indirectness	N	Mean score (SD) on self-rated QoL scale (0 = very poor, 10 = excellent\)	Quality
Carer quality of life (QoL)								
carer-rated QoL	Kirstjanson (2006)	Survey	Serious ¹	N/A ²	Serious ³	141	6.59 (2.27)	LOW
Satisfaction with QoL	Kirstjanson (2006)	Survey	Serious ¹	N/A ²	Serious ³	141	6.35 (2.58)	LOW

Symptom severity experience in patients

Quality asse Number of p	ssment atients								
Example	Studies	Design	Risk of bias	Inconsiste ncy	Indirectness	N	Mean score (SD) on symptom assessment scale (SAS; 0 = no problem, 10=worst problem)	Quality	
Worst experienced symptoms									
Fatigue and tiredness	Kirstjanson (2006)	Survey	Serious ¹	NA ²	Serious ³	174	5.1 (2.9)	LOW	
concentrati on	Kirstjanson (2006)	Survey	Serious ¹	NA ²	Serious ³	174	3.9 (3.1)	LOW	
sleeping	Kirstjanson (2006)	Survey	Serious ¹	NA ²	Serious ³	174	4.1 (3.3)	LOW	

Quality assessment									
Number of pa	tients								
Example	Studies	Design	Risk of bias	Inconsiste ncy	Indirectness	N	Percentage (%) of patients/carers experiencing anxiety and/or depression assessed by Hospital Anxiety Depression Scale (HADS) in patients and General health questionnaire (QHQ) in carers	Quality	
Patient self-reported moderate-to severe experience									
Anxiety	Kirstjanson (2006)	Survey	Serious ¹	N/A ²	Serious ³	174	20%	LOW	
Depression	Kirstjanson (2006)	Survey	Serious ¹	N/A ²	Serious ³	174	30%	LOW	
Carer self-reported moderate-to severe experience									
Anxiety and depression	Kirstjanson (2006)	Survey	Serious ¹	N/A ²	Serious ³	141	19%	LOW	

Incidence of anxiety and depression in patients and carers

 ¹ Serious risk of bias: Methodology not clear, not clear whether all survey material was standardised or validated
² N/A, not applicable, single study
³ Serious indirectness - population was restricted to moderate disease; no advanced or newly diagnosed participants

Parkinson's disease Appendix E