Appendix G: GRADE and CERQual tables

G.1 Dementia diagnosis

G.1.1 Dementia diagnosis

- What are the most effective methods of primary assessment to decide whether a person with suspected dementia should be referred to a dementia service?
- What are the most effective methods of diagnosing dementia and dementia subtypes in specialist dementia diagnostic services?

Please see appendix P

G.1.2 Distinguishing dementia from delirium or delirium with dementia

- What are the most effective methods of differentiating dementia or dementia with delirium from delirium alone?

G.1.2.1 Confusion assessment method (CAM)

No. of studies	Study design	Sample size	Sensitivity (95%CI)	Specificity (95%CI)	Effect size (95%CI)	Risk of bias	Inconsistency	Indirectness	Imprecision	Quality
To distinguish Delirium and Delirium superimposed on Dementia from Dementia										
>5 CAM symptoms										
1 (Cole)	Prospective cohort	262	99.7 (98.5, 100.0)	60.5 (50.6, 70.1)	LR+ 2.53 (1.97, 3.24)	Serious[1]	N/A	Not serious	Serious[2]	Low
					LR- 0.01 (0.00, 0.08)	Serious[1]	N/A	Not serious	Not serious	Moderate
>6 CAM symptoms										
1 (Cole)	Prospective cohort	262	97.6% (94.8, 99.3)	75.5% (66.4, 83.6)	LR+ 3.99 (2.80, 5.70)	Serious[1]	N/A	Not serious	Not serious	Moderate
					LR- 0.03 (0.01, 0.08)	Serious[1]	N/A	Not serious	Not serious	Moderate
To distinguish Delirium from Delirium superimposed on Dementia										
>5 CAM symptoms										
1 (Cole)	Prospective cohort	262	99.6% (98.1, 100)	1.2% (0.00, 6.00)	LR+ 1.01 (0.97, 1.05)	Serious[1]	N/A	Not serious	Not serious	Moderate
					LR- 0.32 (0.01, 15.77)	Serious[1]	N/A	Not serious	Very serious[3]	Very Low
>6 CAM symptoms										
1 (Cole)	Prospective cohort	262	98.4% (95.7, 99.8)	5.00% (0.60, 13.5)	LR+ 1.04 (0.96, 1.1.2)	Serious[1]	N/A	Not serious	Not serious	Moderate
					LR- 0.31 (0.05, 2.15)	Serious[1]	N/A	Not serious	Very serious[3]	Very Low

1. Unclear if people administering CAM were blinded to DSM diagnosis
2. 95% confidence interval for likelihood ratio crosses one end of a defined MID interval – (0.5, 2)

No. of studies	Study design	Sample size	Sensitivity (95%CI)	Specificity (95%CI)	Effect size (95%CI)	Risk of bias	Inconsistency	Indirectness	Imprecision	Quality
3. 95% confidence interval for likelihood ratio crosses both ends of a defined MID interval – (0.5, 2)										

G.1.2.2 Delirium Index (DI)

No. of studies	Study design	Sample size	Sensitivity (95%CI)	Specificity (95%CI)	Effect size (95%CI)	Risk of bias	Inconsistency	Indirectness	Imprecision	Quality
To distinguish Delirium and Delirium superimposed on Dementia from Dementia										
>2 DI symptoms										
1 (Cole)	Prospective cohort	262	89.3% (84.2, 93.5)	29.8% (21.0, 39.4)	LR+ 1.27 (1.10, 1.47)	Serious[1]	N/A	Not serious	Not serious	Moderate
					LR- 0.36 (0.21, 0.61)	Serious[1]	N/A	Not serious	Serious[2]	Low
>3 DI symptoms										
1 (Cole)	Prospective cohort	262	73.2% (66.3, 79.6)	57.4% (47.4, 67.2)	LR+ 1.72 (1.34, 2.21)	Serious[1]	N/A	Not serious	Serious[2]	Low
					LR- 0.47 (0.34, 0.63)	Serious[1]	N/A	Not serious	Serious[2]	Low
>4 DI symptoms										
1 (Cole)	Prospective cohort	262	56.5% (49.0, 63.9)	85.1% (77.3, 91.5)	LR+ 3.80 (2.30, 6.27)	Serious[1]	N/A	Not serious	Not serious	Moderate
					LR- 0.51 (0.42, 0.62)	Serious[1]	N/A	Not serious	Serious[2]	Low
To distinguish Delirium from Delirium superimposed on Dementia										
>2 DI symptoms										
1 (Cole)	Prospective cohort	262	82.4% (69.5, 92.5)	8.6% (4.4, 14.0)	LR+ 0.90 (0.78, 1.05)	Serious[1]	N/A	Not serious	Not serious	Moderate
					LR- 2.04 (0.85, 4.9)	Serious[1]	N/A	Not serious	Serious[2]	Low
>3 DI symptoms										

No. of studies	Study design	Sample size	Sensitivity (95%CI)	Specificity (95%CI)	Effect size (95%CI)	Risk of bias	Inconsistency	Indirectness	Imprecision	Quality
1 (Cole)	Prospective cohort	262	60.0% (44.6, 74.4)	22.7% (15.9,30.3)	LR+ 0.78 (0.59, 1.02)	Serious[1]	N/A	Not serious	Not serious	Moderate
					LR- 1.78 (1.08, 2.90)	Serious[1]	N/A	Not serious	Serious[2]	Low
>4 DI symptoms										
1 (Cole)	Prospective cohort	262	60.9% (52.4, 69.2)	57.5% (42.1, 72.2)	LR+ 1.43 (0.97, 2.11)	Serious[1]	N/A	Not serious	Serious[2]	Low
					LR- 0.68 (0.48, 0.96)	Serious[1]	N/A	Not serious	Serious[2]	Low

1. Unclear if people administering DI were blinded to DSM diagnosis
2. 95% confidence interval for likelihood ratio crosses one end of a defined MID interval – (0.5, 2)

G.1.2.3 Short Portable Mental State Questionnaire (SPMSQ)

No. of studies	Study design	Sample size	Sensitivity (95%CI)	Specificity (95%CI)	Effect size (95%CI)	Risk of bias	Inconsistency	Indirectness	Imprecision	Quality
To distinguish Delirium and Delirium superimposed on Dementia from Dementia										
<3 errors										
1 (Erkinjuntti)	Prospective cohort	70	24.0% (13.1, 36.8)	97.9% (89.8, 100)	LR+ 11.50 (0.71,186.99)	Serious[1]	N/A	Not serious	Serious[2]	Low
					LR- 0.78 (0.66, 0.92)	Serious[1]	N/A	Not serious	Not serious	Moderate
<4 errors										
1 (Erkinjuntti)	Prospective cohort	70	57.4% (43.2, 71.1)	91.3% (77.2, 98.9)	LR+ 6.61 (1.72, 25.41	Serious[1]	N/A	Not serious	Serious[2]	Low
					LR- 0.47 (0.33, 0.67)	Serious[1]	N/A	Not serious	Serious[2]	Low
<5 errors										

No. of studies	Study design	Sample size	Sensitivity (95%CI)	Specificity (95%CI)	Effect size (95%CI)	Risk of bias	Inconsistency	Indirectness	Imprecision	Quality
1 (Erkinjuntti)	Prospective cohort	70	76.6% (63.6, 87.4)	78.3% (59.7, 92.2)	LR+ 3.52 (1.60, 7.77)	Serious[1]	N/A	Not serious	Serious[2]	Low
					LR- 0.30 (0.17,0.52)	Serious[1]	N/A	Not serious	Serious[2]	Low
To distinguish Delirium from Delirium superimposed on Dementia <3 errors										
1 (Erkinjuntti)	Prospective cohort	70	27.4% (15.2, 41.6)	92.9% (67.0, 100)	LR+ 3.83 (0.25, 57.96)	Serious[1]	N/A	Not serious	Serious[2]	Low
					LR- 0.78 (0.59, 1.03)	Serious[1]	N/A	Not serious	Not serious	Moderate
<4 errors										
1 (Erkinjuntti)	Prospective cohort	70	61.0% (45.8, 75.1)	66.7% (28.4, 94.7)	LR+ 1.823 (0.58, 5.82)	Serious[1]	N/A	Not serious	Serious[2]	Low
					LR- 0.59 (0.30, 1.16)	Serious[1]	N/A	Not serious	Serious[2]	Low
<5 errors										
1 (Erkinjuntti)	Prospective cohort	70	82.9% (70.2, 92.7)	66.7% (28.4, 94.7)	LR+ 2.49 (0.80, 7.78)	Serious[1]	N/A	Not serious	Serious[2]	Low
					LR- 0.26 (0.11, 0.62)	Serious[1]	N/A	Not serious	Serious[2]	Low

1. Unclear if people administering SPMSQ were blinded to Dementia Scale diagnosis
2. 95% confidence interval for likelihood ratio crosses one end of a defined MID interval – (0.5, 2)

G.1.2.4 Delirium Rating Scale Revised 98 (DRS-R98)

No. of studies	Study design	Sample size	Sensitivity (95%CI)	Specificity (95%CI)	Effect size (95%CI)	Risk of bias	Inconsistency	Indirectness	Imprecision	Quality
2 studies (Leonard and Trzepacz) but data not comparable so presented separately.										

No. of studies	Study design	Sample size	Sensitivity (95%CI)	Specificity (95%CI)	Effect size (95%CI)	Risk of bias	Inconsistency	Indirectness	Imprecision	Quality
To distinguish Delirium and Delirium superimposed on Dementia from Dementia **Item Severities:** **Sleep-wake cycle disturbance**										
1 (Leonard)	Prospective cohort	144	61.6% (52.5, 70.4)	78.1% (62.5,90.4)	LR+ 2.82 (1.44, 5.51)	Serious[1]	N/A	Not serious	Serious[2]	Low
					LR- 0.49 (0.37, 0.66)	Serious[1]	N/A	Not serious	Serious[2]	Low
Perceptual disturbances and hallucinations										
1 (Leonard)	Prospective cohort	144	26.8% (19.0, 35.3)	93.8% (83.3, 92.2)	LR+ 4.29 (1.10, 17.0)	Serious[1]	N/A	Not serious	Serious[2]	Low
					LR- 0.78 (0.68, 0.90)	Serious[1]	N/A	Not serious	Not serious	Moderate
Delusions										
1 (Leonard)	Prospective cohort	144	15.2% (9.2, 22.4)	90.6% (78.6, 98.0)	LR+ 1.16 (0.51, 5.18)	Serious[1]	N/A	Not serious	Serious[2]	Low
					LR- 0.93 (0.82, 1.07)	Serious[1]	N/A	Not serious	Not serious	Moderate
Lability of affect										
1 (Leonard)	Prospective cohort	144	39.3% (30.5, 48.5)	90.6% (78.6, 98.0)	LR+ 4.19 (1.39, 12.61)	Serious[1]	N/A	Not serious	Serious[2]	Low
					LR- 0.67 (0.56, 0.81)	Serious[1]	N/A	Not serious	Not serious	Moderate
Language										
1 (Leonard)	Prospective cohort	144	30.4% (22.2, 39.1)	90.6% (78.6, 98.0)	LR+ 3.24 (1.06, 9.86)	Serious[1]	N/A	Not serious	Serious[2]	Low
					LR- 0.77 (0.65, 0.91)	Serious[1]	N/A	Not serious	Not serious	Moderate
Thought process abnormalities										

No. of studies	Study design	Sample size	Sensitivity (95%CI)	Specificity (95%CI)	Effect size (95%CI)	Risk of bias	Inconsistency	Indirectness	Imprecision	Quality
1 (Leonard)	Prospective cohort	144	49.1 (39.9, 58.3)	78.1% (62.5, 98.0)	LR+ 2.25 (1.14, 4.44)	Serious[1]	N/A	Not serious	Serious[2]	Low
					LR- 0.65 (0.50, 0.84)	Serious[1]	N/A	Not serious	Not serious	Moderate
Motor agitation										
1 (Leonard)	Prospective cohort	144	38.4% (29.6, 47.5)	84.4% (70.2, 94.5)	LR+ 2.46 (1.06, 5.68)	Serious[1]	N/A	Not serious	Serious[2]	Low
					LR- 0.73 (0.59, 0.90)	Serious[1]	N/A	Not serious	Not serious	Moderate
Motor retardation										
1 (Leonard)	Prospective cohort	144	16.1% (9.9, 23.4)	96.9% (88.8, 99.9)	LR+ 5.14 (0.71, 37.06)	Serious[1]	N/A	Not serious	Serious[2]	Low
					LR- 0.87 (0.78, 0.96)	Serious[1]	N/A	Not serious	Not serious	Moderate
Orientation										
1 (Leonard)	Prospective cohort	144	45.5% (36.3, 54.8)	78.1% (62.5, 90.4)	LR+ 2.08 (1.05, 4.13)	Serious[1]	N/A	Not serious	Serious[2]	Low
					LR- 0.70 (0.54, 0.90)	Serious[1]	N/A	Not serious	Not serious	Moderate
Attention										
1 (Leonard)	Prospective cohort	144	75.9% (67.6, 83.3)	68.8% (52.0, 83.3)	LR+ 2.43 (1.44, 4.10)	Serious[1]	N/A	Not serious	Serious[2]	Low
					LR- 0.35 (0.23, 0.52)	Serious[1]	N/A	Not serious	Serious[2]	Low
Short-term memory										
1 (Leonard)	Prospective cohort	144	65.2% (56.2, 73.7)	40.6% (24.5, 57.8%	LR+ 1.10 (0.80, 1.51)	Serious[1]	N/A	Not serious	Serious[2]	Low

No. of studies	Study design	Sample size	Sensitivity (95%CI)	Specificity (95%CI)	Effect size (95%CI)	Risk of bias	Inconsistency	Indirectness	Imprecision	Quality
					LR- 0.86 (0.53, 1.40)	Serious[1]	N/A	Not serious	Not serious	Moderate
Long-term memory										
1 (Leonard)	Prospective cohort	144	42.0% (33.0, 51.2)	68.8% (52.0, 83.3)	LR+ 1.34 (0.77, 2.35)	Serious[1]	N/A	Not serious	Serious[2]	Low
					LR- 0.84 (0.64, 1.12)	Serious[1]	N/A	Not serious	Not serious	Moderate
Visuospatial ability										
1 (Leonard)	Prospective cohort	144	64.3% (55.2, 72.9)	40.6% (24.5, 57.8)	LR+ 1.08 (0.77, 2.35)	Serious[1]	N/A	Not serious	Serious[2]	Low
					LR- 0.88 (0.54, 1.43)	Serious[1]	N/A	Not serious	Not serious	Moderate
Temporal onset of symptoms										
1 (Leonard)	Prospective cohort	144	64.3% (55.2, 72.9)	87.5% (74.2, 96.4)	LR+ 5.14 (2.04, 13.00)	Serious[1]	N/A	Not serious	Not serious	Moderate
					LR- 0.41 (0.31, 0.54)	Serious[1]	N/A	Not serious	Serious[2]	Low
Fluctuation in symptom severity										
1 (Leonard)	Prospective cohort	144	17.0% (10.6, 24.4)	71.9% (55.4, 85.8)	LR+ 0.60 (0.30, 1.20)	Serious[1]	N/A	Not serious	Serious[2]	Low
					LR- 1.16 (0.92, 1.46)	Serious[1]	N/A	Not serious	Not serious	Moderate
Physical disorder										
1 (Leonard)	Prospective cohort	144	87.5% (80.8, 92.9)	65.6% (48.6, 80.8)	LR+ 2.55 (1.57, 4.13)	Serious[1]	N/A	Not serious	Serious[2]	Low
					LR- 0.19 (0.11, 0.33)	Serious[1]	N/A	Not serious	Not serious	Moderate
To distinguish Delirium from Delirium superimposed on Dementia										

No. of studies	Study design	Sample size	Sensitivity (95%CI)	Specificity (95%CI)	Effect size (95%CI)	Risk of bias	Inconsistency	Indirectness	Imprecision	Quality
Item Severities: **Sleep-wake cycle disturbance**										
1 (Leonard)	Prospective cohort	112	74.0% (61.1, 85.1)	46.8% (34.6, 59.2)	LR+ 1.39 (1.05, 1.85)	Serious[1]	N/A	Not serious	Not serious	Moderate
					LR- 0.56 (0.33, 0.95)	Serious[1]	N/A	Not serious	Serious[2]	Low
Perceptual disturbances and hallucinations										
1 (Leonard)	Prospective cohort	112	32.0% (119.9, 45.4)	77.4% (63.3, 86.8)	LR+ 1.42 (0.77, 2.62)	Serious[1]	N/A	Not serious	Serious[2]	Low
					LR- 0.88 (0.70, 1.11)	Serious[1]	N/A	Not serious	Not serious	Moderate
Lability of affect										
1 (Leonard)	Prospective cohort	112	48.0% (34.4, 61.7)	67.7% (55.7, 78.7)	LR+ 1.49 (0.94, 2.36)	Serious[1]	N/A	Not serious	Serious[2]	Low
					LR- 0.77 (0.56, 1.05)	Serious[1]	N/A	Not serious	Not serious	Moderate
Language										
1 (Leonard)	Prospective cohort	112	40.0% (27.7, 53.8)	77.4% (66.3, 86.8)	LR+ 1.77 (1.00, 3.14)	Serious[1]	N/A	Not serious	Serious[2]	Low
					LR- 0.78 (0.60, 1.01)	Serious[1]	N/A	Not serious	Not serious	Moderate
Thought process abnormalities										
1 (Leonard)	Prospective cohort	112	64.0% (50.4, 76.6)	61.3% (49.0, 72.9)	LR+ 1.65 (1.14, 2.41)	Serious[1]	N/A	Not serious	Serious[2]	Low
					LR- 0.59 (0.39, 0.89)	Serious[1]	N/A	Not serious	Serious[2]	Low
Motor agitation										

No. of studies	Study design	Sample size	Sensitivity (95%CI)	Specificity (95%CI)	Effect size (95%CI)	Risk of bias	Inconsistency	Indirectness	Imprecision	Quality
1 (Leonard)	Prospective cohort	112	20.0% (10.2, 32.0)	87.1% (77.8, 94.2)	LR+ 1.55 (0.66, 3.63)	Serious[1]	N/A	Not serious	Serious[2]	Low
					LR- 0.92 (0.78, 1.10)	Serious[1]	N/A	Not serious	Not serious	Moderate
Orientation										
1 (Leonard)	Prospective cohort	112	38.0% (25.2, 51.7)	48.4% (36.1, 60.7)	LR+ 0.74 (0.48, 1.13)	Serious[1]	N/A	Not serious	Serious[2]	Low
					LR- 1.28 (0.92, 1.79)	Serious[1]	N/A	Not serious	Not serious	Moderate
Attention										
1 (Leonard)	Prospective cohort	112	80% (68.0, 89.8)	24.7% (17.1, 39.1)	LR+ 1.10 (0.90, 1.36)	Serious[1]	N/A	Not serious	Not serious	Moderate
					LR- 0.73 (0.37 (1.45)	Serious[1]	N/A	Not serious	Serious[2]	Low
Temporal onset of symptoms										
1 (Leonard)	Prospective cohort	112	78.0% (65.7, 88.2)	46.8% (34.6, 59.2)	LR+ 1.47 (1.11, 1.93)	Serious[1]	N/A	Not serious	Not serious	Moderate
					LR- 0.47 (0.26, 0.85)	Serious[1]	N/A	Not serious	Serious[2]	Low
Physical disorder										
1 (Leonard)	Prospective cohort	112	92.0% (83.1, 97.7)	16.1% (8.2, 26.2)	LR+ 1.10 (0.96, 1.26)	Serious[1]	N/A	Not serious	Not serious	Moderate
					LR- 0.50 (0.17, 1.49)	Serious[1]	N/A	Not serious	Serious[2]	Low
2nd study Cut off score 17.75 DRS-98 Total										
1	Case-control	37	97.8% (89.3, 100)	82.1% (59.1, 96.7)	LR+ 5.48 (1.78, 16.88)	Serious[3]	N/A	Serious[4]	Serious[2]	Very Low

No. of studies	Study design	Sample size	Sensitivity (95%CI)	Specificity (95%CI)	Effect size (95%CI)	Risk of bias	Inconsistency	Indirectness	Imprecision	Quality
(Trzepacz)					LR- 0.03 (0.00, 0.42)	Serious[3]	N/A	Serious[4]	Not serious	Low
Cut off score 21.50 DRS-98 Total										
1 (Trzepacz)	Case-control	37	90.9% (76.2, 98.8)	92.3% (73.5, 99.8)	LR+ 11.82 (1.79, 78.05)	Serious[3]	N/A	Serious[4]	Serious[2]	Very Low
					LR- 0.09 (0.03, 0.37)	Serious[3]	N/A	Serious[4]	Not serious	Low
Cut off score 22.50 DRS-98 Total										
1 (Trzepacz)	Case-control	37	89.1% (73.9, 98.1)	96.4 % (82.7, 100)	LR+ 24.96 (1.64, 380.98)	Serious[3]	N/A	Serious[4]	Serious[2]	Very Low
					LR- 0.11 (0.04, 0.37)	Serious[3]	N/A	Serious[4]	Not serious	Low
2nd study Cut off score 15.25 DRS-98 Severity										
1 (Trzepacz)	Case-control	37	97.8% (89.3, 100)	75.9% (50.3, 93.0)	LR+ 3.91 (1.58, 9.72)	Serious[3]	N/A	Serious[4]	Serious[2]	Very Low
					LR- 0.03 (0.00, 0.46)	Serious[3]	N/A	Serious[4]	Not serious	Low
Cut off score 17.00 DRS-98 Severity										
1 (Trzepacz)	Case-control	37	86.4% (69.6, 97.0)	92.3% (73.5, 99.8)	LR+ 11.23 (1.70, 74.35)	Serious[3]	N/A	Serious[4]	Serious[2]	Very Low
					LR- 0.15 (0.05, 0.43)	Serious[3]	N/A	Serious[4]	Not serious	Low

1. Unclear if people administering DRS-R98 were blinded to DSM IV diagnosis.
2. 95% confidence interval for likelihood ratio crosses one end of a defined MID interval – (0.5, 2)
3. Patients selected for dementia or delirium at baseline and research assistant screened patients for suitability before DRS-R98 was carried out.
4. Patients not randomly/ consecutively selected and then diagnosed as in scope

G.1.2.5 Cognitive Test for Delirium (CTD)

No. of studies	Study design	Sample size	Sensitivity (95%CI)	Specificity (95%CI)	Effect size (95%CI)	Risk of bias	Inconsistency	Indirectness	Imprecision	Quality
To distinguish Delirium and Delirium superimposed on Dementia from Dementia <4 CTD SSF points										
1 (Meagher)	Prospective cohort	100	63.8% (53.0, 73.9)	85.0% (66.9, 96.6)	LR+ 4.25 (1.48, 12.21)	Serious[1]	N/A	Not serious	Serious[2]	Low
					LR- 0.43 (0.30, 0.60)	Serious[1]	N/A	Not serious	Serious[2]	Low
To distinguish Delirium from Delirium superimposed on Dementia <4 CTD SSF points										
1 (Meagher)	Prospective cohort	100	65.9% (48.9, 78.8)	37.5% (23.4, 52.8)	LR+ 1.04 (0.74, 1.45)	Serious[1]	N/A	Not serious	Not serious	Moderate
					LR- 0.93 (0.52, 1.67)	Serious[1]	N/A	Not serious	Not serious	Moderate

1. Unclear if people administering CTD were blinded to DSM diagnosis
2. 95% confidence interval for likelihood ratio crosses one end of a defined MID interval – (0.5, 2)

G.1.2.6 Observational Scale of Level of Arousal (OSLA) and OSLA combined with the Attention Test

No. of studies	Study design	Sample size	Sensitivity (95%CI)	Specificity (95%CI)	Effect size (95%CI)	Risk of bias	Inconsistency	Indirectness	Imprecision	Quality
To distinguish Delirium and Delirium superimposed on Dementia from No Delirium (Dementia and No dementia or delirium) >4 OSLA										
1 (Richardson)	Prospective cohort	114	84.6% (73.7, 93.0)	82.3% (71.9, 90.6)	LR+ 4.70 (2. 76, 8.25)	Serious[1]	N/A	Serious[3]	Not serious	Low
					LR- 0.19 (0.09, 0.36)	Serious[1]	N/A	Serious[3]	Not serious	Low
To distinguish Delirium and Delirium superimposed on Dementia from No Delirium (Dementia and No dementia delirium) >9 Combination of OSLA and Attention Test										

No. of studies	Study design	Sample size	Sensitivity (95%CI)	Specificity (95%CI)	Effect size (95%CI)	Risk of bias	Inconsistency	Indirectness	Imprecision	Quality
1 (Richardson)	Prospective cohort	114	84.6% (73.7, 93.0)	96.8% 91.2, 99.6)	LR+ 26.23 (6.68, 103.050)	Serious[2]	N/A	Serious[3]	Not serious	Low
					LR- 0.16 (0.08, 0.30)	Serious[2]	N/A	Serious[3]	Not serious	Low
To distinguish Delirium superimposed on Dementia from Dementia >4 OSLA										
1 (Richardson)	Prospective cohort	59	74.2% (57.7, 87.7)	96.4% (87.2, 99.9)	LR+ 20.77 (3.00, 143.96)	Serious[1]	N/A	Serious[3]	Not serious	Low
					LR- 0.27 (0.15, 0.49)	Serious[1]	N/A	Serious[3]	Not serious	Low
To distinguish Delirium superimposed on Dementia from Dementia >9 Combination of OSLA and Attention Test										
1 (Richardson)	Prospective cohort	59	93.5% (82.2, 99.2)	92.9% (81.0, 99.1)	LR+ 13.10 (3.43, 49.95)	Serious[2]	N/A	Serious[3]	Not serious	Low
					LR- 0.069 (0.02, 0.27)	Serious[2]	N/A	Serious[3]	Not serious	Low

1. Unclear whether people administering the index test were blinded to reference diagnosis.
2. Unclear whether people administering the index test were blinded to reference diagnosis and use of an optimised threshold for the attention test.
3. Participants were > 70 years old as part of the inclusion criteria
4. 95% confidence interval for likelihood ratio crosses one end of a defined MID interval – (0.5, 2)

G.1.3 Case finding for people at high risk of dementia

- What are the most effective methods of case finding for people at high risk of dementia?

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
New diagnoses of dementia and MCI together among stage 1 participants (with general estimating equation applied to account for clustering)							
1 (van den Dungen 2016)	Not serious	N/A	Serious[1]	Very serious[3]	647	RR 1.33 (0.70, 2.07)*	Very low
New diagnoses of dementia and MCI together among stage 2 participants (adjusted for Activities of Daily Living, ADL, and instrumental ADL dependency)							
1 (van den Dungen 2016)	Not serious	N/A	Serious[1]	Very serious[3]	145	RR 1.07 (0.60, 1.62)*	Very low
Mental Health Elderly (MH5) at baseline (range 0-100)							
1 (van den Dungen 2016)	Not serious	N/A	Not serious	Serious[2]	124	MD 1.59 (-5.04, 8.22)	Moderate
Mental Health Elderly (MH5) at 6 months (range 0-100)							
1 (van den Dungen 2016)	Not serious	N/A	Not serious	Serious[2]	124	MD 2.11 (-3.31, 7.53)	Moderate
Mental Health Elderly (MH5) at 12 months (range 0-100)							
1 (van den Dungen 2016)	Not serious	N/A	Not serious	Serious[2]	124	MD 0.21 (-6.35, 6.77)	Moderate
Mental health close relative (GHQ12) at baseline (range 0-12, higher scores indicate worse health)							
1 (van den Dungen 2016)	Not serious	N/A	Not serious	Serious[2]	104	MD -0.08 (-1.06, 0.90)	Moderate
Mental health close relative (GHQ12) at 6 months (range 0-12, higher scores indicate worse health)							
1 (van den Dungen 2016)	Not serious	N/A	Not serious	Serious[2]	104	MD -0.30 (-1.19, 0.59)	Moderate
Mental health close relative (GHQ12) at 12 months (range 0-12, higher scores indicate worse health)							
1 (van den Dungen 2016)	Not serious	N/A	Not serious	Serious[2]	104	MD -0.33 (-1.30, 0.64)	Moderate
Quality of life elderly (EQ5D) at baseline (range -0.33-1)							

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
1 (van den Dungen 2016)	Not serious	N/A	Not serious	Serious[2]	124	MD -0.03 (-0.10, 0.04)	Moderate
Quality of life elderly (EQ5D) at 6 months (range -0.33-1)							
1 (van den Dungen 2016)	Not serious	N/A	Not serious	Serious[2]	124	MD -0.02 (-0.09, 0.05)	Moderate
Quality of life elderly (EQ5D) at 12 months (range -0.33-1)							
1 (van den Dungen 2016)	Not serious	N/A	Not serious	Serious[2]	124	MD -0.03 (-0.10, 0.04)	Moderate
Quality of life elderly (QoL-AD) at baseline (range 13-52)							
1 (van den Dungen 2016)	Not serious	N/A	Not serious	Serious[2]	124	MD -0.23 (-2.06, 1.60)	Moderate
Quality of life elderly (QoL-AD) at 6 months (range 13-52)							
1 (van den Dungen 2016)	Not serious	N/A	Not serious	Serious[2]	124	MD -0.61 (-2.31, 1.09)	Moderate
Quality of life elderly (QoL-AD) at 12 months (range 13-52)							
1 (van den Dungen 2016)	Not serious	N/A	Not serious	Serious[2]	124	MD -0.85 (-2.46, 0.76)	Moderate
Quality of life close relative (EQ5D) at baseline (range -0.33-1)							
1 (van den Dungen 2016)	Not serious	N/A	Not serious	Serious[2]	104	MD -0.04 (-0.11, 0.03)	Moderate
Quality of life close relative (EQ5D) at 6 months (range -0.33-1)							
1 (van den Dungen 2016)	Not serious	N/A	Not serious	Serious[2]	104	MD -0.01 (-0.07, 0.05)	Moderate
Quality of life close relative (EQ5D) at 12 months (range -0.33-1)							
1 (van den Dungen 2016)	Not serious	N/A	Not serious	Serious[2]	104	MD -0.03 (-0.09, 0.03)	Moderate
Sense of competence to provide care, close relative (SSQC) at baseline (range 0-35)							
1 (van den Dungen 2016)	Not serious	N/A	Not serious	Serious[2]	104	MD -0.86 (-2.70, 0.98)	Moderate

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
Sense of competence to provide care, close relative (SSQC) at 6 months (range 0-35)							
1 (van den Dungen 2016)	Not serious	N/A	Not serious	Serious[2]	104	MD -0.88 (-2.58, 0.82)	Moderate
Sense of competence to provide care, close relative (SSQC) at 12 months (range 0-35)							
1 (van den Dungen 2016)	Not serious	N/A	Not serious	Serious[2]	104	MD -0.79 (-2.49, 0.91)	Moderate

1. Data is for MCI and dementia groups combined. MCI is out of guideline scope.
2. Non-significant result.
3. 95% CI crosses 2 lines of a defined MID interval

*RR calculated from OR reported in paper.

G.2 Involving people with dementia in decision about care

G.2.1 Barriers and facilitators to involvement in decision making for people living with dementia

- What barriers and facilitators have an impact on involving people living with dementia in decisions about their present and future care?
- What barriers and facilitators have an impact on how people living with dementia can make use of advance planning?

G.2.1.1 Barriers to decision making

Studies	Study design	Description	Methodological limitations	Relevance	Coherence	Adequacy	Confidence
Patient level - Denial of problem							
3 (Goodman, Livingston, Poppe)	Focus groups, interviews	If the person with dementia is unreconciled to the severity of their needs, this is a barrier to accepting care. The main barrier to advance planning on the part of the people with dementia and carers was difficulty for some people with dementia or carers to accept the diagnosis.	Not serious	High	High	High	High
Patient level - Rejection of help							
1 (Livingston)	Focus groups, interviews	People will often reject help, either because they feel they do not need it or because accepting help would involve psychologically acknowledging the severity of their problems.	Not serious	High	High	High	High
Patient level – Deference to authority							
1 (Goodman)	Interviews	Having dementia combined with living in a care home meant the older people often accepted that staff and visiting healthcare professionals would make decisions on their behalf.	Very serious[1]	High	Moderate[2]	Moderate[3]	Very low
1 (Goodman)	Interviews	Knowing that they had dementia affected confidence in expressing opinions, self-esteem and whether they thought their views were worth listening to.	Very serious[1]	High	Moderate[2]	Moderate[3]	Very low
Patient level – Poor relationship with formal or informal carers							

Studies	Study design	Description	Methodological limitations	Relevance	Coherence	Adequacy	Confidence
1 (Goodman)	Interviews	If the person with dementia has a poor relationship with the carer(s), this could be a barrier to expressing a wish regarding care.	Very serious[1]	High	Moderate[2]	Moderate[3]	Very low
Patient level – one partner more dominant							
1 Dening (2017)	Semi-structured interviews	Often there was one partner more dominant in decision-making.	Not serious	High	Moderate[2]	High	Moderate
Professional – Not recognising problems							
1 (Livingston)	Focus groups, interviews	Healthcare professionals may not recognise people need additional assistance to be involved in decision-making particularly when people are not open about difficulties they are having.	Not serious	High	High	High	High
Professional – Late diagnosis							
1 (Livingston)	Focus groups, interviews	If the diagnosis of dementia is delayed, this can make it difficult for all the necessary advance discussions to be had before capacity issues start to occur.	Not serious	High	High	High	High
Professional – Timing and quantity of information given							
2 (Livingston, Lord)	Focus groups, interviews	Feelings of guilt and distress for carers were often exacerbated by a perceived lack of support and information.	Not serious	High	High	High	High
Professional - Confidentiality and data protection							
1 (Livingston)	Focus groups, interviews	Carers felt they could not get the necessary information to help support decision-making because of confidentiality issues.	Not serious	High	High	High	High
Professional – Bureaucracy and rigidity (sticking to protocols)							
1 (Livingston)	Focus groups, interviews	People felt discussions were not sufficiently individualised due to a reliance on following pre-specified protocols.	Not serious	High	High	High	High
Carer – Role conflict							

Studies	Study design	Description	Methodological limitations	Relevance	Coherence	Adequacy	Confidence
2 (Livingston, Lord)	Focus groups, interviews	Many carers reported the decision was against the care recipient's wishes, and signalled a major carer role transition. Carers report a shift in the dynamic to a "mother/child" type relationship. They struggled with being expected to relinquish their caregiver role and that friends and family perceived the dyadic relationship to be over.	Not serious	High	High	High	High
Carer – Relationship to person living with dementia							
1 (Samsi)	Interviews	Friend carers often felt they were less able to make decisions on behalf of individuals than family carers.	Serious[4]	High	High	Moderate[3]	Low
Carer – Carer guilt							
2 (Livingston, Lord)	Focus groups, interviews	Feelings of anguish and guilt over decisions made. Journey towards a decision was directed by a mixture of fatigue and a lack of obvious or available alternatives. Feelings of guilt and failure were particularly strong for people obliged to cope alone.	Not serious	High	High	High	High
Carer – Family conflict							
2 (Livingston, Samsi)	Focus groups, interviews	When the person with dementia was involved in decision-making, they usually expressed reluctance to move to a care home. This often led the carer either to delay the decision or exclude the person with dementia from decision-making.	Not serious	High	High	High	High
Carer – Rigidity of system							
1 (Livingston)	Focus groups, interviews	People felt that once a decision was reached, it was then difficult to change this decision if circumstances changed, and this led to a reluctance to make initial decisions.	Not serious	High	High	High	High
Carer – Cultural issues							
2 (Lord, Mackenzie)	Interviews	Cultural issues may place a particular strain on decision-making around future places of care. In South	Not serious	Moderate[5]	High	High	Moderate

Studies	Study design	Description	Methodological limitations	Relevance	Coherence	Adequacy	Confidence
		Asian communities, there may be a tendency to want to protect the person with dementia from ridicule by keeping them away from other people.					
Structural – Inability to plan							
2 (Lord, Poppe)	Interviews	Struggle with knowing when to seek care home placement due to dementia being unpredictable and wait lists of institutions. Some patients find discussing the future difficult without knowing what the future will bring.	Not serious	High	High	High	High
1. Theme only identified in studies at high risk of bias 2. Theme does not consistently emerge from all relevant studies 3. Insufficient data to develop a full understanding of the phenomenon of interest 4. Theme only identified in studies at moderate or high risk of bias 5. Unclear how the groups included in this study generalise to the population at large							

G.2.1.2 Facilitators for decision making

Studies	Study design	Description	Methodological limitations	Relevance	Coherence	Adequacy	Confidence
Patient – Reconceptualisation and adjustment							
1 (Livingston)	Focus groups, interviews	Re-conceptualisation of services as optimising independence. Allowing services to develop slowly.	Not serious	High	High	High	High
Professional – Providing practical support							
2 (Livingston, Lord)	Focus groups, interviews	Suggesting interventions to facilitate agreement, or structured approaches to decision making. Collaboration with staff helped carers with decision-making, and this was facilitated by a trusted healthcare professional who consulted them and advocated effectively	Not serious	High	High	High	High

Studies	Study design	Description	Methodological limitations	Relevance	Coherence	Adequacy	Confidence
1 (Livingston)	Focus groups, interviews	Providing high-quality information in a timely fashion.	Not serious	High	High	High	High
Professional – Initiating conversations							
1 (Lord)	Focus groups, interviews	Carers felt that clinician's raising these discussions helped them with decision-making	Not serious	High	High	High	High
Professional – Legal and financial issues							
1 (Livingston)	Focus groups, interviews	Ensuring the patient is asked to give permission for information to be given to carers. Access to legal and financial advice.	Not serious	High	High	High	High
Professional – Structured tools							
1 (Poppe)	Interviews	Open-ended, structured tools may be useful to guide discussions around advance planning. Staff who had not yet conducted any advance care planning discussions themselves were unsure how to initiate the discussion with those people with dementia who had not raised the issue themselves, but saw the tool as a potential way of facilitating this.	Serious[1]	High	High	Moderate[2]	Low
Carer - Participation							
1 (Livingston)	Focus groups, interviews	Carer accompanying patient on visits to healthcare professionals. Posing a question to the person at the "right" time, gauging when their relative was likely to be most engaged in conversation, and presenting a limited number of options.	Not serious	High	High	High	High
Carer – Shared decision-making							
2 (Livingston, Lord)	Focus groups, interviews	Carers found it helpful to hear the perspectives of other members of the family or professionals when making decision on behalf of the person with dementia – they felt it "gave permission" to make decisions.	Not serious	High	High	High	High

Studies	Study design	Description	Methodological limitations	Relevance	Coherence	Adequacy	Confidence
Carer – Family cohesion							
2 (Livingston, Lord)	Focus groups, interviews	Not feeling that different members of the family are pulling in different directions. Carers often sought reassurance after decision making from other family members.	Not serious	High	High	High	High
Structural – Social support							
1 (Livingston)	Focus groups, interviews	Extended family, voluntary and community networks.	Not serious	High	High	High	High
Intervention – Talking Mats							
1 (Murphy)	Interviews	Discussing care was facilitated by using Talking Mats. Talking Mats helped the participants with dementia to be aware of what their family members were doing for them, and were seen an enjoyable activity which improved communication between the person with dementia and his/her family.	Serious[1]	High	High	Moderate[2]	Low

1. Theme only identified in studies at moderate or high risk of bias
2. Insufficient data to develop a full understanding of the phenomenon of interest

G.2.1.3 Issues identified in Huntington's disease

Studies	Study design	Description	Methodological limitations	Relevance	Coherence	Adequacy	Confidence
Barrier/facilitator – Information provision							
1 (Bisson)	Interviews	Some confusion was apparent among people with Huntington's disease regarding what advance decisions and powers of attorney are, not least the difference between advance decisions and euthanasia.	Not serious	Moderate[1]	High	Moderate[2]	Low
1 (Bisson)	Interviews	Easy-to-follow, consistent verbal and written information was desired, which should be Huntington's disease specific.	Not serious	Moderate[1]	High	Moderate[2]	Low

Studies	Study design	Description	Methodological limitations	Relevance	Coherence	Adequacy	Confidence
1 (Bisson)	Interviews	Involvement in the care pathway was a positive experience for the majority.	Not serious	Moderate[1]	High	Moderate[2]	Low
Facilitator – Therapeutic relationships							
1 (Bisson)	Interviews	A facilitator for advance planning is having an established therapeutic relationship with an expert in Huntington's disease. Personal qualities such as being approachable, caring and sensitive with good communication skills were felt to be important. Participants also recommended the additional offer of home visits by a Huntington's disease Association Advisor.	Not serious	Moderate[1]	High	Moderate[2]	Low
Facilitator - Early introduction to advance decisions							
1 (Bisson)	Interviews	Opinions of patients with Huntington's disease were different to professionals. Professionals were reluctant to approach service users too early, particularly asymptomatic individuals with the altered Huntington's disease gene, for fear of causing distress.	Not serious	Moderate[1]	High	Moderate[2]	Low
1 (Bisson)	Interviews	The earlier discussions regarding advance decisions are introduced the better, subject to checking personal circumstances and support, to allow consideration of them before individuals develop symptoms or their symptoms worsen.	Not serious	Moderate[1]	High	Moderate[2]	Low
1 (Bisson)	Interviews	It was considered important to have a minimum 2-week "cool off" period between an initial meeting and advance decision completion. The duration should be flexible allowing for as many sessions required to reach a decision.	Not serious	Moderate[1]	High	Moderate[2]	Low
Facilitator - Advance decision forms							
1 (Bisson)	Interviews	The main issues that people believed should be on the form were: life-saving treatments, percutaneous endoscopic gastrostomy feeding, location of future care,	Not serious	Moderate[1]	High	Moderate[2]	Low

Studies	Study design	Description	Methodological limitations	Relevance	Coherence	Adequacy	Confidence
		capacity assessment, witness details and a distribution list. A summary sheet for patient files and checklists for education, completion and review were considered important. Participants suggested adding statements concerning organ donation and whether independent legal advice had been received.					
Facilitator – Power of attorney							
1 (Bisson)	Interviews	The power of attorney information was considered to be too detailed to be included on the advance decision form. Therefore, a single booklet containing all the information was recommended.	Not serious	Moderate[1]	High	Moderate[2]	Low

1. Some people in the study were positive for the Huntington's disease gene but did not yet have a diagnosis of Huntington's disease
2. Insufficient data to develop a full understanding of the phenomenon of interest

G.3 Care planning, review and co-ordination

G.3.1 Health and social care co-ordination

Review questions

- What are the most effective methods of care planning, focussing upon improving outcomes for people with dementia and their carers?
- How should health and social care be co-ordinated for people living with dementia?

G.3.1.1 CERQual tables

Themes identified for the self-management intervention for people living with dementia and their carers

Studies	Study design	Description	Methodological limitations	Relevance	Coherence	Adequacy	Confidence
Theme: The program training was enjoyable							
1 (Faith 2015)	Focus groups, interviews	Although people living with dementia said that they could not recall all of the activities, they had enjoyed the program.	Serious[1]	High	High	Moderate[3]	Low
Theme: The participants felt empowered							
2 (Faith 2015, Moore 2011)	Focus groups, interviews	The training program encouraged people living with dementia to continue with their hobbies and goals (Faith 2015). Access to a budget provided a sense of empowerment (Moore 2011).	Serious[1]	High	High	High	Moderate
Theme: Caregivers felt burdened and people living with dementia felt disempowered							
1 (Toms 2015)	Semi-structured interviews	The caregivers felt responsible and burdened. This left the person with dementia feeling disempowered.	Not serious	High	Moderate[2]	Moderate[3]	Low
Theme: Support groups were considered valuable							

Studies	Study design	Description	Methodological limitations	Relevance	Coherence	Adequacy	Confidence
1 (Toms 2015)	Semi-structured interviews	Peer support, such as support groups, was considered valuable by participants.	Not serious	High	Moderate[2]	Moderate[3]	Low
Theme: Caregivers and people with dementia questioned what would happen once time-limited support ended							
1 (Toms 2015)	Semi-structured interviews	Additional support, such as a support group, was available, but these were often time-limited, which led both caregivers and people with dementia to the question of what happened when such support ended.	Not serious	High	Moderate[2]	Moderate[3]	Low
Theme: There was a lack of support							
1 (Toms 2015)	Semi-structured interviews	People living with dementia and their caregivers felt that there was a lack of support.	Not serious	High	Moderate[2]	Moderate[3]	Low
Theme: Respondents thought that professional support was important for effective self-management							
1 (Toms 2015)	Semi-structured interviews	Respondents thought that professional support was important for effective self-management, and valued this resource. They thought that this help was necessary because not everything could be self-managed within the family.	Not serious	High	Moderate[2]	Moderate[3]	Low
Theme: Many respondents were unsure how to access the services and reported finding them limited and poorly integrated							
1 (Toms 2015)	Semi-structured interviews	Many respondents were unsure how to access the services that were available, and reported finding them limited and poorly integrated. This made it harder to self-manage the condition.	Not serious	High	Moderate[2]	Moderate[3]	Low
Theme: Some people living with dementia used practical aids to support their memory							
1 (Toms 2015)	Semi-structured interviews	Some people living with dementia used practical aids to support their memory.	Not serious	High	Moderate[2]	Moderate[3]	Low
Theme: What was most pertinent to carers was the diminished ability of the person living with dementia to complete daily tasks							

Studies	Study design	Description	Methodological limitations	Relevance	Coherence	Adequacy	Confidence
1 (Toms 2015)	Semi-structured interviews	What was most pertinent to carers was the diminished ability of the person living with dementia to complete daily tasks.	Not serious	High	Moderate[2]	Moderate[3]	Low
Theme: The approach of normalising difficulties was evident in many interviews							
1 (Toms 2015)	Semi-structured interviews	The approach of normalising difficulties was evident in many interviews.	Not serious	High	Moderate[2]	Moderate[3]	Low
Theme: People living with dementia and their carers endured hardship without showing their feelings or complaining							
1 (Toms 2015)	Semi-structured interviews	A sense of stoicism, often expressed when respondents gave their ideas about self-management, was evident in many interviews, and this seemed to be a form of psychological management.	Not serious	High	Moderate[2]	Moderate[3]	Low
Theme: People with dementia were uncertain about the future. This led to lack of confidence and a diminished belief that they could self-manage							
1 (Toms 2015)	Semi-structured interviews	Some people with dementia discussed losing confidence. It was implied that this loss of confidence could diminish people's belief that they could self-manage. In some cases, this loss of confidence seemed to relate to uncertainty about the future and how the illness would progress	Not serious	High	Moderate[2]	Moderate[3]	Low
Theme: Diaphragmatic breathing was relaxing							
1 (Faith 2015)	Focus groups, semi-structured interviews	Participants found the relaxation activity of diaphragmatic breathing relaxing	Serious[1]	High	High	Moderate[3]	Low
Theme: Funding for respite was useful for carers							
1 (Moore 2011)	Interviews	Funding for respite was useful for carers	Serious[1]	High	Moderate[2]	Moderate[3]	Very low
Theme: Finding personal assistants was difficult							

Studies	Study design	Description	Methodological limitations	Relevance	Coherence	Adequacy	Confidence
1 (Moore 2011)	Interviews	Finding suitable individuals to become personal assistants was difficult for some people	Serious[1]	High	Moderate[2]	Moderate[3]	Very low
Theme: When suitable individuals became personal assistants, there were positive results							
1 (Moore 2011)	Interviews	When suitable individuals became personal assistants, there were positive results	Serious[1]	High	Moderate[2]	Moderate[3]	Very low

1. Theme only identified in studies at high risk of bias.
2. This theme conflicts with another. The difference may be partially, although not completely explained by the fact that participants in Moore 2011 had access to a budget and those in Toms 2015 did not.
3. Only a limited amount of evidence to support this finding.

Themes identified for outcome-focussed/needs-led care vs standard care

Studies	Study design	Description	Methodological limitations	Relevance	Coherence	Adequacy	Confidence
Theme: Standard care: Familial carers often feel not able to cope							
1 (Gethin-Jones 2014)	Semi-structured interviews	The most common concern of familial carers is the feeling of not being able to cope	Not serious	High	High	Moderate[1]	Moderate
Theme: Standard care: Carers felt isolated							
1 (Gethin-Jones 2014)	Semi-structured interviews	The sense of isolation expressed by the participants came over very strongly. This isolation appeared to come from their sense that they were on the outside with little control because the care was planned by the other professionals. Family carers felt that they were isolated as they had all the responsibility and in their eyes and potentially all the blame when things went wrong.	Not serious	High	High	Moderate[1]	Moderate
Theme: Outcome-focussed care: Carers' self-reported well-being improved after the outcome-focused intervention had been implemented							
2 (Gethin-Jones 2014,	Semi-structured interviews	There was an improvement in the carers' self-reported subjective well-being, after the outcome-focused homecare intervention had been implemented.	Not serious	High	High	High	High

Studies	Study design	Description	Methodological limitations	Relevance	Coherence	Adequacy	Confidence
Rothera 2008)							
Theme: Outcome-focussed care: Carers felt the subjective well-being of their relative had improved after the outcome-focused care intervention							
1 (Gethin-Jones 2014)	Semi-structured interviews	All the carers felt the subjective well-being of their relative had improved after the six month outcome-focused care intervention.	Not serious	High	High	Moderate[1]	Moderate
1. Only a limited amount of evidence to support this finding.							

Themes identified for community-based case management

Studies	Study design	Description	Methodological limitations	Relevance	Coherence	Adequacy	Confidence
Theme: Meeting health and social care professionals at home was more relaxing and less stressful							
1 (Gibson 2007)	Interviews	Meeting health and social care professionals at home was more relaxing and less stressful compared to using the memory service.	Not serious	High	High	Moderate[1]	Moderate
Theme: Being at home facilitated communication							
1 (Gibson 2007)	Interviews	Being at home facilitated communication with health and social care professionals.	Not serious	High	High	Moderate[1]	Moderate
Theme: The case manager was good at identifying needs and providing the right support							
1 (Iliffe 2014)	Interviews	The case manager was good at identifying needs and providing the right support.	Not serious	High	High	Moderate[1]	Moderate
Theme: Carers expected case managers to provide information about dementia and services							
1 (Iliffe 2014)	Interviews	Carers expected case managers to provide information about dementia and services.	Not serious	High	High	Moderate[1]	Moderate
Theme: Case managers should be proactive in asking carers and people living with dementia if they feel they need assistance							
1 (Iliffe 2014)	Interviews	Case managers should be proactive in asking carers and people living with dementia if they feel they need assistance. This is because participants frequently expressed a reluctance to initiate contact with the case	Not serious	High	High	Moderate[1]	Moderate

Studies	Study design	Description	Methodological limitations	Relevance	Coherence	Adequacy	Confidence
		manager, which undermines the concept that they could ask for help when needed.					
Theme: A common reason why people living with dementia and their carers do not initiate contact with case managers is because they do not associate case managers with assisting with day-to-day issues							
1 (Iliffe 2014)	Interviews	A common reason why people living with dementia and their carers do not initiate contact with case managers is because they associate case managers with assisting with 'major' problems such as arranging residential care homes. They do not associate case managers with assisting with day-to-day issues.	Not serious	High	High	Moderate[1]	Moderate
Theme: People living with dementia and their carers preferred to have their case manager based at their GP's surgery							
1 (Iliffe 2014)	Interviews	People living with dementia and their carers preferred to have their case manager based at their GP's surgery. This is because there was the perception that their GP's surgery would then be a 'one-stop shop'. In addition, having the case manager at the GP's surgery provided an additional opportunity to talk to the case manager while visiting the GP's surgery.	Not serious	High	High	Moderate[1]	Moderate
Theme: Appointments at clinics were more anxiety provoking compared to home appointments							
1 (Gibson 2007)	Interviews	For some, exposure to others at more severe stages of the illness within the clinic was a potent contributor towards anxiety, illustrating what could be expected as the disease progresses. Appointments at home removed this exposure.	Not serious	High	High	Moderate[1]	Moderate
Theme: Nurses as case managers were perceived as providing a more direct link to the GP for advice and support							
1 (Iliffe 2014)	Interviews	From the perspectives of some people living with dementia and their carers, nurses as case managers were perceived as providing a more direct link to the GP for advice and support for comorbidities and minor ailments.	Not serious	High	Moderate[2]	Moderate[1]	Low
Theme: A direct link to the GP was not a priority because they preferred their case manager to have expertise in social services							

Studies	Study design	Description	Methodological limitations	Relevance	Coherence	Adequacy	Confidence
1 (Iliffe 2014)	Interviews	From the perspectives of some people living with dementia and their carers, a direct link to the GP was not a priority because they preferred their case manager to have expertise in social services. The inference is that they would prefer a social worker to be the case manager.	Not serious	High	Moderate[2]	Moderate[1]	Low
Theme: People living with dementia and their carers emphasised interpersonal skills							
1 (Iliffe 2014)	Interviews	People living with dementia and their carers emphasised interpersonal skills such as empathy.	Not serious	High	High	Moderate[1]	Moderate
Theme: Case management made access to services easier							
1 (Iliffe 2014)	Interviews	Case management made access to services easier including GPs, benefit checks and links to other services.	Not serious	High	High	Moderate[1]	Moderate
Theme: Case managers should respond as quickly as possible to questions							
1 (Iliffe 2014)	Interviews	Case managers should respond as quickly as possible to questions from people living with dementia or their carers.	Not serious	High	High	Moderate[1]	Moderate
Theme: The idea of background support was valued by people living with dementia and their carers							
1 (Iliffe 2014)	Interviews	A key aspect of case management valued by people living with dementia and their carers was the idea of background support that could easily be called on at a time of need.	Not serious	High	High	Moderate[1]	Moderate
Theme: There needed to be time and opportunities to develop a deeper relationship.							
1 (Iliffe 2014)	Interviews	For people living with dementia and their carers to feel comfortable about contacting the case manager in the event of difficulties, there needed to be time and opportunities to develop a deeper relationship.	Not serious	High	High	Moderate[1]	Moderate
Theme: Face-to-face contact was preferred							
1 (Iliffe 2014)	Interviews	Face-to-face and telephone contact were both considered acceptable, although face-to-face contact	Not serious	High	High	Moderate[1]	Moderate

Studies	Study design	Description	Methodological limitations	Relevance	Coherence	Adequacy	Confidence
		was often preferred as it facilitated relationship building better than telephone contact.					
Theme: Some people living with dementia and their carers do not mind contact by telephone							
1 (Iliffe 2014)	Interviews	Some people living with dementia and their carers appreciate the service that case managers provide and also appreciate how hard they work. Therefore, they do not mind contact by telephone.	Not serious	High	High	Moderate[1]	Moderate
Theme: Case managers should explain what support they can provide							
1 (Iliffe 2014)	Interviews	Case managers should explain to carers, and where appropriate to people living with dementia, what support they can provide.	Not serious	High	High	Moderate[1]	Moderate
Theme: Participants found case management more useful than dementia advisors							
1 (Iliffe 2014)	Interviews	Participants found case management more useful than dementia advisors. This is because case management offers continuity of care but dementia advisors do not.	Not serious	High	High	Moderate[1]	Moderate

1. Only a limited amount of evidence to support this finding.
2. This finding conflicts with another. Among people living with dementia and there carers, opinion is divided.

Themes identified for memory-clinic case management

Studies	Study design	Description	Methodological limitations	Relevance	Coherence	Adequacy	Confidence
Theme: The memory service was well received							
1 (Hean 2011)	Interviews	The memory service was well received.	Very serious[1,2]	High	High	Moderate[3]	Very low
Theme: People living with dementia experienced an increase in their quality of life							
1 (Sonola 2013)	Focus groups, survey	People living with dementia generally experienced an increase in their quality of life.	Serious[2]	High	High	Moderate[3]	Low
Theme: Familial carers' stress scores improved or remained stable							

Studies	Study design	Description	Methodological limitations	Relevance	Coherence	Adequacy	Confidence
1 (Sonola 2013)	Focus groups, survey	Familial carers' stress scores improved or remained stable for all the carers measured.	Serious[2]	High	High	Moderate[3]	Low
Theme: There was difficulty and effort in accessing treatment							
1 (Gibson 2007)	Interviews	There was difficulty and effort in accessing treatment	Not serious	High	High	Moderate[3]	Moderate
Theme: For memory services that do not have post-diagnostic support, participants expressed feelings of abandonment							
1 (Kelly 2016)	Semi-structured interviews	For memory services that do not have post-diagnostic support, many participants expressed feelings of abandonment or 'being sent away' by professionals on receipt of diagnosis.	Not serious	High	High	Moderate[3]	Moderate
Theme: For memory services that do have post-diagnostic support, participants explained the value of having support as soon after diagnosis as possible							
1 (Kelly 2016)	Semi-structured interviews	For memory services that do have post-diagnostic support, people with dementia and their carers explained the value of having support as soon after diagnosis as possible and the importance of skilled, knowledgeable, sensitive project workers to deliver support.	Not serious	High	High	Moderate[3]	Moderate
Theme: Carers frequently reported positively on the help received from the project workers with claiming benefits							
1 (Kelly 2016)	Semi-structured interviews	Carers frequently reported positively on the help received from the project workers with claiming benefits.	Not serious	High	High	Moderate[3]	Moderate
Theme: Carers spoke of receiving support with arranging Power of Attorney							
1 (Kelly 2016)	Semi-structured interviews	Carers spoke of receiving support with arranging Power of Attorney and valued the input from project workers in negotiating the process.	Not serious	High	High	Moderate[3]	Moderate
Theme: Participants found the information they received useful							
1 (Kelly 2016)	Semi-structured interviews	Family members and one person newly diagnosed with dementia found the information they received (books and leaflets) along with general advice useful.	Not serious	High	High	Moderate[3]	Moderate

Studies	Study design	Description	Methodological limitations	Relevance	Coherence	Adequacy	Confidence
Theme: Exposure to others at more severe stages of the illness within the clinic was a potent contributor towards anxiety							
1 (Gibson 2007)	Interviews	For some, exposure to others at more severe stages of the illness within the clinic was a potent contributor towards anxiety, illustrating what could be expected as the disease progresses. Appointments at home removed this exposure.	Not serious	High	High	Moderate[3]	Moderate
Theme: The coordination of care was valued							
2 (Hean 2011, Sonola 2013)	Interviews, focus groups, survey	The coordination of care was valued.	Not serious	High	High	High	High
Theme: The service made carers and people living with dementia feel supported and reassured							
2 (Hean 2011, Sonola 2013)	Interviews, focus groups, survey	The service and nature of the staff made carers and people living with dementia feel supported and reassured. (Having a named person to contact in times of crisis, and the security that they would not left to manage alone.)	Not serious	High	High	High	High
Theme: The language used was not quite right							
1 (Hean 2011)	Interviews	The language used was not quite right.	Very serious[1,2]	High	High	Moderate[3]	Very low
Theme: People living with dementia felt pressure of time because the psychiatrist was busy							
1 (Hean 2011)	Interviews	People living with dementia felt pressure of time because the psychiatrist was busy.	Very serious[1,2]	High	High	Moderate[3]	Very low
Theme: Some found it difficult to get to the right people and get the answers needed							
1 (Hean 2011)	Interviews	Some found it difficult to get to the right people and get the answers needed.	Very serious[1,2]	High	High	Moderate[3]	Very low
Theme: There were accounts of receiving insufficient information							

Studies	Study design	Description	Methodological limitations	Relevance	Coherence	Adequacy	Confidence
1 (Kelly 2016)	Semi-structured interviews	There were accounts of receiving no information, or insufficient or inappropriate information following diagnosis.	Not serious	High	High	Moderate[3]	Moderate
Theme: Some carers expressed discomfort with some of the information they received							
1 (Kelly 2016)	Semi-structured interviews	Some carers expressed discomfort with some of the information they received. Some felt that it was too much to face too soon. Many participants stated that a 'one size fits all' approach was not what they wanted.	Not serious	High	High	Moderate[3]	Moderate
Theme: Participants valued information that was delivered on a one-to-one basis and targeted to individual needs and wishes							
1 (Kelly 2016)	Semi-structured interviews	Participants valued that information was delivered by the project workers on a one-to-one basis and specifically targeted to individual needs and wishes.	Not serious	High	High	Moderate[3]	Moderate
Theme: People living with dementia and their carers liked seeing the same person throughout treatment							
2 (Hean 2011, Willis 2011)	Interviews, semi-structured interviews	People living with dementia and their carers liked seeing the same person throughout treatment.	Not serious	High	High	High	High
Theme: People living with dementia and their carers recognised the one stop shop aspect of the memory service.							
1 (Willis 2011)	Semi-structured interviews	Convenience. People living with dementia and their carers recognised the one stop shop aspect of the memory service. Ten participants described the memory service as a central point of access to all necessary services.	Serious[2]	High	High	Moderate[3]	Low
Theme: People living with dementia and their carers thought that home visits were very good							
1 (Hean 2011)	Interviews	People living with dementia and their carers thought that home visits were very good.	Very serious[1,2]	High	High	Moderate[3]	Very low
Theme: People living with dementia and their carers valued transport that was arranged by case managers/project workers.							
1 (Kelly 2016)	Semi-structured interviews	People living with dementia and their carers valued transport that was arranged by case managers/project workers.	Not serious	High	High	Moderate[3]	High

Studies	Study design	Description	Methodological limitations	Relevance	Coherence	Adequacy	Confidence
Theme: Care management does not promote advance care planning							
1 (Kelly 2016)	Semi-structured interviews	Care management does not promote advance care planning.	Not serious	High	High	Moderate[3]	Moderate
Theme: Memory service post-diagnostic support when individualised and one-to-one, causes people with dementia to re-engage							
1 (Kelly 2016)	Semi-structured interviews	Memory service post-diagnostic support when individualised and one-to-one, causes people with dementia to re-engage socially or with old hobbies.	Not serious	High	High	Moderate[3]	Moderate

1. Method of recruitment not mentioned. Recruitment numbers not clarified.
2. Theme only identified in studies at high risk of bias.
3. Only a limited amount of evidence to support this finding.

Themes identified for Daisy Chain: a commercial person-centred dementia service that seems to have some elements of case management

Studies	Study design	Description	Methodological limitations	Relevance	Coherence	Adequacy	Confidence
Theme: The person-centred community-based dementia service was well received							
1 (Gladman 2007)	Observation and semi-structured interviews	The person-centred community-based dementia service was well received.	Not serious	Moderate[1]	High	Moderate[2]	Low
Theme: The person-centred community-based dementia service provides a personalised service							
1 (Gladman 2007)	Observation and semi-structured interviews	The person-centred community-based dementia service provides a personalised service.	Not serious	Moderate[1]	High	Moderate[2]	Low
Theme: The person-centred community-based dementia service helped carers to cope							

Studies	Study design	Description	Methodological limitations	Relevance	Coherence	Adequacy	Confidence
1 (Gladman 2007)	Observation and semi-structured interviews	The person-centred community-based dementia service helped carers to cope.	Not serious	Moderate[1]	High	Moderate[2]	Low
Theme: The person-centred community-based dementia service kept the people living with dementia and their accommodation clean							
1 (Gladman 2007)	Observation and semi-structured interviews	The person-centred community-based dementia service kept the people living with dementia and their accommodation clean.	Not serious	Moderate[1]	High	Moderate[2]	Low
Theme: The person-centred community-based dementia service enabled people living with dementia to stay at home							
1 (Gladman 2007)	Observation and semi-structured interviews	The person-centred community-based dementia service enabled people living with dementia to stay at home.	Not serious	Moderate[1]	High	Moderate[2]	Low
Theme: The person-centred community-based dementia service had good communication							
1 (Gladman 2007)	Observation and semi-structured interviews	The person-centred community-based dementia service had good communication.	Not serious	Moderate[1]	High	Moderate[2]	Low
Theme: There is a ‘right time’ for someone living with dementia to move to a residential care home							
1 (Gladman 2007)	Observation and semi-structured interviews	There is a ‘right time’ for someone living with dementia to move to a residential care home.	Not serious	Moderate[1]	High	Moderate[2]	Low
Theme: Some carers would prefer the person living with dementia to remain in their own home							

Studies	Study design	Description	Methodological limitations	Relevance	Coherence	Adequacy	Confidence
1 (Gladman 2007)	Observation and semi-structured interviews	Some carers would prefer the person living with dementia to remain in their own home.	Not serious	Moderate[1]	High	Moderate[2]	Low
Theme: There are sometimes differences of opinion							
1 (Gladman 2007)	Observation and semi-structured interviews	There are sometimes differences of opinion between people living with dementia, paid carers and familial carers.	Not serious	Moderate[1]	High	Moderate[2]	Low

1. Full details of what is contained in the intervention are unclear.
2. Only a limited amount of evidence to support this finding.

Themes identified for non-specified case management style(s) in predominantly remote and rural areas in Scotland

Studies	Study design	Description	Methodological limitations	Relevance	Coherence	Adequacy	Confidence
Theme: Carers said they required more help							
1 (Innes 2014)	Semi-structured interviews	Carers generally expressed satisfaction with support received but said they required more help	Serious[1]	High	High	Low[2]	Very low
Theme: The lack of alternative options sometimes led to provision of no support at all							
1 (Innes 2014)	Semi-structured interviews	The lack of alternative options sometimes led to provision of no support at all.	Serious[1]	High	High	Low[2]	Very low
Theme: Poor coordination of services							
1 (Gorska 2013, Innes 2014)	Semi-structured interviews	Poor coordination of services. The participants particularly emphasized poor communication between existing services, which results in unsatisfactory case management and delays in service provision. The need	Not serious	High	High	High	High

Studies	Study design	Description	Methodological limitations	Relevance	Coherence	Adequacy	Confidence
		for a single point of access to information and service coordination was expressed as a means to manage these challenges and to facilitate more efficient and effective service delivery. Participant reports also highlighted inconsistencies in care provision and suggested the need for well-defined care pathways.					
Theme: Some experienced lack of continuity of care							
1 (Gorska 2013, Innes 2014)	Semi-structured interviews	Some experienced lack of continuity of care. This can lead to poor communication and is confusing.	Not serious	High	High	High	High
Theme: Lack of mental stimulation							
1 (Gorska 2013)	Semi-structured interviews	Lack of mental stimulation.	Not serious	High	High	Low[2]	Low
Theme: Some people living with dementia do not want to make use of day centres							
1 (Innes 2014)	Semi-structured interviews	Some people living with dementia do not want to make use of day centres.	Serious[1]	High	High	Low[2]	Very low
Theme: Some GPs have a specific interest in dementia and this improves communication							
1 (Innes 2014)	Semi-structured interviews	One interviewee pointed out that some GPs have a specific interest in dementia and this improves communication.	Serious[1]	High	High	Low[2]	Very low
Theme: There were high satisfaction levels with the support received from the Community Mental Health Team							
1 (Innes 2014)	Semi-structured interviews	There were high satisfaction levels with the support received from the Community Mental Health Team.	Serious[1]	High	High	High	Moderate
Theme: Participants discussed the importance of staff building a rapport with the person with dementia							

Studies	Study design	Description	Methodological limitations	Relevance	Coherence	Adequacy	Confidence
1 (Innes 2014)	Semi-structured interviews	Participants discussed the importance of staff building a rapport with the person with dementia. This facilitates communication.	Serious[1]	High	High	Low[2]	Very low
Theme: When it was available, a carers' group was appreciated							
1 (Innes 2014)	Semi-structured interviews	When it was available, a carers' group (caregiver support) was appreciated.	Serious[1]	High	High	Low[2]	Very low
Theme: Practical support was important to carers who received help from services regularly							
1 (Innes 2014)	Semi-structured interviews	Practical support was important to most carers who received help from private or voluntary services regularly. Carers perceived this type of support as an opportunity to take a respite from caregiving responsibilities. Many used the respite time to rest, run errands which required getting out, or to attend carers meetings.	Serious[1]	High	High	Low[2]	Very low
Theme: Other sources of post-diagnostic support were from family, friends, and neighbours							
1 (Innes 2014)	Semi-structured interviews	Other sources of post-diagnostic support were from family, friends, and neighbours.	Serious[1]	High	High	High	Moderate
Theme: Some carers have difficulty leaving their relative with someone else							
1 (Innes 2014)	Semi-structured interviews	Some carers have difficulty leaving their relative with someone else.	Serious[1]	High	High	Low[2]	Very low
Theme: Information was not always in a format appropriate for the person with dementia or carers							
1 (Innes 2014)	Semi-structured interviews	Information was not always in a format appropriate for the person with dementia or carers.	Serious[1]	High	High	High	Moderate
Theme: Participants preferred a direct approach when receiving information with the opportunity to ask questions							

Studies	Study design	Description	Methodological limitations	Relevance	Coherence	Adequacy	Confidence
1 (Innes 2014)	Semi-structured interviews	The way information was delivered was important. Participants preferred a direct approach with the opportunity to ask questions.	Serious[1]	High	High	High	Moderate
Theme: Care managers should be proactive in anticipating the needs of people living with dementia and their carers							
1 (Innes 2014)	Semi-structured interviews	Care managers should be proactive in anticipating the needs of people living with dementia and their carers and provide relevant information.	Serious[1]	High	High	Low[2]	Very low

1. Methods of recruitment are not described.
2. Very limited amount of evidence to support this finding.

Themes identified for case management in residential care homes

Studies	Study design	Description	Methodological limitations	Relevance	Coherence	Adequacy	Confidence
Theme: The need for activities, interaction and outings was the most prevalent theme overall							
1 (Popham 2012)	Focus groups, interviews	The need for activities, interaction and outings was the most prevalent theme overall.	Not serious	High	High	Moderate[1]	Moderate
Theme: Participants valued freedom to carry out normal everyday activities and domestic chores							
1 (Popham 2012)	Focus groups, interviews	Participants spoke about having the freedom to be able to carry out normal everyday activities and domestic chores.	Not serious	High	High	Moderate[1]	Moderate
Theme: Rooms with views were highly valued							
1 (Popham 2012)	Focus groups, interviews	Rooms with views were highly valued.	Not serious	High	High	Moderate[1]	Moderate

1. Only a limited amount of evidence to support this finding.

Case planning – the Adaption-Coping Model

Studies	Study design	Description	Methodological limitations	Relevance	Coherence	Adequacy	Confidence
Family carers also valued having the opportunity to learn more about dementia and see other people in the same situation.							
1 (Brooker 2017)	Focus group interviews	It enabled some carers to gain a broader perspective on their own experiences, and facilitate adjustment. By seeing how their relatives were treated at the Meeting Centre and responded to the interactions, some carers were able to reflect on the difficulties faced in their everyday lives.	Serious[1]	High	High	High	Moderate
Participants liked the warmth and friendliness of the staff							
1 (Brooker 2017)	Focus group interviews	Participants liked the warmth and friendliness of the staff. It gave them confidence.	Serious[1]	High	High	High	Moderate
The Meeting Centre provides a supportive space for feelings to be aired							
1 (Brooker 2017)	Focus group interviews	Some carers felt that they were unable to share their true feelings or experiences with family members for fear of judgement, and again the Meeting Centre provides a supportive space for those feelings to be aired	Serious[1]	High	High	High	Moderate
The experience enabled some people to reflect upon their own emotional adjustment							
1 (Brooker 2017)	Focus group interviews	The experience enabled some people to reflect upon their own emotional adjustment	Serious[1]	High	High	High	Moderate
The planned activity provided a useful structure							
1 (Brooker 2017)	Focus group interviews	The planned activity provided a useful structure	Serious[1]	High	High	High	Moderate
The participants felt that they were not alone							
1 (Brooker 2017)	Focus group interviews	The participants felt that they were not alone	Serious[1]	High	High	High	Moderate
Carers were able to get a different perspective							

Studies	Study design	Description	Methodological limitations	Relevance	Coherence	Adequacy	Confidence
1 (Brooker 2017)	Focus group interviews	Seeing other people in similar situations and getting outside perceptions helped one carer to reassess how he views his wife's situation	Serious[1]	High	High	High	Moderate
Attendance was good							
1 (Brooker 2017)	Focus group interviews	The participants enjoyed attending and therefore the attendance was good	Serious[1]	High	High	High	Moderate

1. Theme only identified in one study at moderate risk of bias

Case planning – Rotherham Carers Resilience Service

Studies	Study design	Description	Methodological limitations	Relevance	Coherence	Adequacy	Confidence
Carer – Often people suggested that they felt unsure and extremely anxious about the person they were caring for							
1 Dayson (2016)	Interviews	Often people suggested that they felt unsure and extremely anxious about the person they were caring for	Serious[1]	High	High	High	Moderate
Carer – Carers felt that the service provided them with a great deal of reassurance, both in practical terms but also emotional							
1 Dayson (2016)	Interviews	Carers felt that the service provided them with a great deal of reassurance, both in practical terms but also emotional	Serious[1]	High	High	High	Moderate
Carer – The relief people felt moving forwards							
1 Dayson (2016)	Interviews	Understanding that the situation will change in the future, beneficiaries of the service described how their knowledge of the service helped them to feel more positive about the future	Serious[1]	High	High	High	Moderate
Carer – Participants felt supported							
1 Dayson (2016)	Interviews	People now felt 'in the system', and felt reassured knowing where they could go for support should anything occur in the future.	Serious[1]	High	High	High	Moderate
Carer – Carers reported that the knowledge and experience of the staff was key							

Studies	Study design	Description	Methodological limitations	Relevance	Coherence	Adequacy	Confidence
1 Dayson (2016)	Interviews	Carers were reassured by the expertise of the staff.	Serious[1]	High	High	High	Moderate
Carer – Carers found that they had benefited from the information provided							
1 Dayson (2016)	Interviews	This is because they had learnt something new or had been reassured that what they were experiencing was not an isolated case	Serious[1]	High	High	High	Moderate
Carer – Carers received practical assistance							
1 Dayson (2016)	Interviews	Examples of help ranged from assessments of homes, recommending alarms and safety devices, through to benefits advice and information about community transport and the provision of a home based support service, whereby a care support worker can come to sit with someone for support and reassurance whilst their carer/partner is away	Serious[1]	High	High	High	Moderate

1. Theme only identified in one study at moderate risk of bias

Coordination – for people living with dementia who have comorbidity

Studies	Study design	Description	Methodological limitations	Relevance	Coherence	Adequacy	Confidence
Family members were often proactive in facilitating continuity and negotiating access to services for their relatives with dementia.							
1 Bunn (2017)	Semi-structured interviews	This included acting as an advocate for their family member with dementia, noticing when something was wrong and seeking help	Serious[1]	High	High	High	Moderate
Family members were often proactive in helping clinicians make treatment decisions, such as whether to thrombolyse a PLWD after a stroke.							
1 Bunn (2017)	Semi-structured interviews	Family carers also had a significant role in coordinating their relative's care, navigating healthcare systems and facilitating continuity of care; for example, managing appointments, organising transport, keeping records of test results and medication	Serious[1]	High	High	High	Moderate
Family members were often proactive in actively transferring information between HCPs and different services							

Studies	Study design	Description	Methodological limitations	Relevance	Coherence	Adequacy	Confidence
1 Bunn (2017)	Semi-structured interviews	Family members were often proactive in actively transferring information between HCPs and different services	Serious[1]	High	High	High	Moderate
The availability of a family carer to act as a proxy, and provide consent, information and post-discharge support impacted on a PLWD's access to care.							
1 Bunn (2017)	Semi-structured interviews	HCPs recognised that PLWD who lived alone, or did not have support from a family carer or advocate, were particularly vulnerable and may have poorer access to care	Serious[1]	High	High	High	Moderate
Although HCPs in our study valued the role family carers played, there was little formal recognition of the carers' role, and no systems for negotiating how or when carers' views could be incorporated into care planning.							
1 Bunn (2017)	Semi-structured interviews	This was reflected in the many examples provided by their interviews where carers felt undervalued or excluded from decision-making about their relative's care.	Serious[1]	High	High	High	Moderate
There were many challenges for family carers.							
1 Bunn (2017)	Semi-structured interviews	These included difficulty in understanding how health systems worked and who to contact, their own health problems, emotional and practical challenges of changing roles	Serious[1]	High	High	High	Moderate
Living at a distance and/or with work and family commitments that made taking on responsibilities for day-to-day care difficult.							
1 Bunn (2017)	Semi-structured interviews	Caring at a distance may be particularly problematic for carers of PLWD as it is difficult for them to offer support or to monitor adherence to medication over the phone.	Serious[1]	High	High	High	Moderate
Support from social networks, such as extended family, friends and religious groups, and from third sector providers were clearly important to PLWD and their carers.							
1 Bunn (2017)	Semi-structured interviews	Support from social networks, such as extended family, friends and religious groups, and from third sector providers were clearly important to PLWD and their carers.	Serious[1]	High	High	High	Moderate
Formal support from health and social care was often seen as inadequate.							

Studies	Study design	Description	Methodological limitations	Relevance	Coherence	Adequacy	Confidence
1 Bunn (2017)	Semi-structured interviews	Formal support from health and social care was often seen as inadequate.	Serious[1]	High	High	High	Moderate
PLWD and family carers valued continuity, in terms of relationships with practitioners but also in terms of encounters that factored in the impact of dementia, that built on earlier conversations and appointments and that included people with dementia and their carers in decision-making.							
1 Bunn (2017)	Semi-structured interviews	Many PLWD and carers reported positive relationships with their GPs and recognised the role that GPs played in coordinating care.	Serious[1]	High	High	High	Moderate
How PLWD managed their care, for example, either independently, in tandem with a family carer or with external health and social care support, was linked to where they were on the dementia trajectory.							
1 Bunn (2017)	Semi-structured interviews	Some people with early stage dementia were still able to self-manage their care. As the dementia got worse, the PLWD's ability to self-manage declined and responsibility moved, either partly or totally, from the PLWD to a carer. These transitions often happened when strategies to facilitate self-management, for example, memory aids, diaries and dosette boxes, ceased to be effective	Serious[1]	High	High	High	Moderate
Current infrastructure did not support the sharing of information across different specialities.							
1 Bunn (2017)	Semi-structured interviews	Current infrastructure did not support the sharing of information across different specialities.	Serious[1]	High	High	High	Moderate
For many participants, their comorbid health condition predated the diagnosis of dementia.							
1 Bunn (2017)	Semi-structured interviews	Despite this, there appeared to be inadequate consideration by some services of the implications of a diagnosis of dementia on the management of existing conditions.	Serious[1]	High	High	High	Moderate

1. Theme only identified in one study at moderate risk of bias

G.3.1.2 GRADE tables

Care coordination/management using a protocol/action plan (that involves educating the carers) and meeting every 3 months vs usual care

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Usual care	Summary of results	
Care recipient's quality of life (DQoL): overall perception on quality of life (higher values favour intervention)									
1 (Jansen 2011)	RCT	Not serious	Not serious	N/A	Serious[1]	43	38	MD 0.40 (-0.50, 1.30)	Moderate
Caregiver sense of competence: consequences of involvement in care (higher values favour intervention)									
1 (Jansen 2011)	RCT	Not serious	Not serious	N/A	Serious[1]	43	38	MD 0.10 (-0.19, 0.39)	Moderate
Caregiver's sense of competence: satisfaction with the older adult (higher values favour intervention)									
1 (Jansen 2011)	RCT	Not serious	Not serious	N/A	Serious[1]	43	38	MD 0.50 (-1.63, 2.63)	Moderate
Caregiver's quality of life (SF-36): mental component summary (higher values favour intervention)									
1 (Jansen 2011)	RCT	Not serious	Not serious	N/A	Serious[1]	43	38	MD -2.50 (-6.82, 1.82)	Moderate
Caregiver's quality of life (SF-36): physical component summary (higher values favour intervention)									
1 (Jansen 2011)	RCT	Not serious	Not serious	N/A	Serious[1]	43	38	MD 2.00 (-2.20, 6.20)	Moderate
Caregiver's depressive symptoms (higher values favour control)									
1 (Jansen 2011)	RCT	Not serious	Not serious	N/A	Serious[1]	43	38	MD 0.60 (-0.25, 1.45)	Moderate
Caregiver's burden (higher values favour control)									
1 (Jansen 2011)	RCT	Not serious	Not serious	N/A	Serious[1]	43	38	MD 0.30 (-0.55, 1.15)	Moderate
Caregiver sense of competence: satisfaction with one's own performance (higher values favour intervention)									
1 (Jansen 2011)	RCT	Not serious	Not serious	N/A	Serious[1]	43	38	MD 0.10 (-0.02, 0.22)	Moderate

1. Non-significant result

Care coordination/management using a protocol/action plan (that involves educating the carers) and peer support group meetings every 2 months vs usual care

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Usual care	Summary of results	
Percentage of people living with dementia who had been admitted to long-term institutional care by the end of the study (higher values favour control)									

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Usual care	Summary of results	
1 (Eloniemi-Sulkava 2009)	RCT	Serious[1]	Not serious	N/A	Serious[2]	63	62	MD -4.10 (-21.69, 13.49)	Low

1. No blinding, attrition rates are not mentioned, not all clinically relevant outcomes were reported (e.g. caregiver burden, ADLs, NPI)
2. Non-significant result

Care coordination/management with monthly follow-up calls and visits every 3 months

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Usual care	Summary of results	
Carer outcome: depression (values greater than 1 favour control)									
1 (Schoenmakers 2010)	RCT	Serious[1]	Not serious	N/A	Serious[2]	23	23	OR 0.16 (0.03, 0.86)	Low
Carer outcome: burden (values greater than 1 favour control)									
1 (Schoenmakers 2010)	RCT	Serious[1]	Not serious	N/A	Serious[2]	23	23	OR 0.09 (0.01, 1.10)	Low
Carer outcome: anxiety (values greater than 1 favour control)									
1 (Schoenmakers 2010)	RCT	Serious[1]	Not serious	N/A	Very serious[3]	23	23	OR 0.30 (0.05, 2.30)	Very low
Carer outcome: emotional coping (values greater than 1 favour control)									
1 (Schoenmakers 2010)	RCT	Serious[1]	Not serious	N/A	Serious[2]	23	23	OR 0.10 (0.01, 1.20)	Low
Carer outcome: supporting coping (values greater than 1 favour control)									
1 (Schoenmakers 2010)	RCT	Serious[1]	Not serious	N/A	Serious[2]	23	23	OR 0.20 (0.03, 1.10)	Low
Carer outcome: problem solving (values greater than 1 favour control)									

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Usual care	Summary of results	
1 (Schoenmakers 2010)	RCT	Serious[1]	Not serious	N/A	Very serious[3]	23	23	OR 0.20 (0.03, 1.60)	Very low
Person living with dementia outcome: frailty (values greater than 1 favour control)									
1 (Schoenmakers 2010)	RCT	Serious[1]	Not serious	N/A	Very serious[3]	23	23	OR 0.20 (0.03, 1.30)	Very low
Person living with dementia outcome: IADL dependency (values greater than 1 favour control)									
1 (Schoenmakers 2010)	RCT	Serious[1]	Not serious	N/A	Serious[2]	23	23	OR 0.20 (0.02, 1.10)	Low
Person living with dementia outcome: incontinence (values greater than 1 favour control)									
1 (Schoenmakers 2010)	RCT	Serious[1]	Not serious	N/A	Serious[2]	23	23	OR 0.20 (0.03, 1.04)	Low
Person living with dementia outcome: disruptive behaviour (values greater than 1 favour control)									
1 (Schoenmakers 2010)	RCT	Serious[1]	Not serious	N/A	Very serious[3]	23	23	OR 0.10 (0.03, 1.90)	Very low
Person living with dementia outcome: mood swings (values greater than 1 favour control)									
1 (Schoenmakers 2010)	RCT	Serious[1]	Not serious	N/A	Very serious[3]	23	23	OR 0.10 (0.01, 1.20)	Very low
Person living with dementia outcome: neurovegetative disturbances (values greater than 1 favour control)									
1 (Schoenmakers 2010)	RCT	Serious[1]	Not serious	N/A	Serious[2]	23	23	OR 0.10 (0.01, 0.98)	Low
Person living with dementia outcome: psychotic features (values greater than 1 favour control)									
1 (Schoenmakers 2010)	RCT	Serious[1]	Not serious	N/A	Very serious[3]	23	23	OR 0.10 (0.01, 1.40)	Very low

1. The number of events in either group are not reported. Therefore, only the relative difference is reported, not the absolute difference.

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Usual care	Summary of results	
2. 95% CI crosses one line of a defined MID interval 3. 95% CI crosses two lines of a defined MID interval									

Care coordination/management using a protocol/action plan (that involves educating the carers) and monthly meetings vs usual care

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Usual care	Summary of results	
Care recipient depression in dementia (higher values favour control)									
1 (Callahan 2006)	RCT	Not serious	Not serious	N/A	Serious[1]	65	49	MD -0.20 (-1.75, 1.35)	Moderate
Mean number of hospital admissions (higher values favour control)									
2 (Bass 2003, Bass 2015)	RCT	Serious[2,3,4,5]	Not serious	Not serious	Serious[1]	298	187	MD 0.01 (-0.15, 0.17)	Low
Percentage of participants who had emergency department visits (higher values favour control)									
1 (Bass 2015)	RCT	Serious[2,5]	Not serious	N/A	Serious[9]	206	122	RR 0.95 (0.74, 1.21)	Low
Mean number of emergency department visits (higher values favour control)									
2 (Bass 2003, Bass 2015)	RCT	Serious[2,3,4,5]	Not serious	Not serious	Serious[1]	298	187	MD -0.13 (-0.38, 0.11)	Low
Percentage institutionalised by the end of the study (cumulative long-term institutionalisation) (higher values favour control)									
2 (Eloniemi-Sulkava 2001, Fortinsky 2009)	RCT	Serious[2,3,5]	Not serious	Serious[6]	Very serious[10]	107	77	RR 0.73 (0.34, 1.59)	Very low
Percentage of people living with dementia who were placed by the end of the study (higher values favour control)									
1 (Chu 2000)	RCT	Serious[2,3]	Not serious	N/A	Not serious	33	36	OR 0.35 (0.17, 0.74)	Moderate
Unmet needs (change from 6 months to 12 months) (higher values favour control)									
2 (Bass 2013, Bass 2014)	RCT	Serious[2,3,7]	Not serious	Not serious	Serious[9]	421	259	SMD -0.28 (-0.44, -0.13)	Low
Care recipient embarrassment - low six-month T2 cognitive impairment (0 to 3) (higher values favour control)									
1 (Bass 2014)	RCT	Serious[2,3,7]	Not serious	N/A	Not serious	122	72	MD 0.20 (0.03, 0.37)	Moderate
Care recipient embarrassment - high six-month T2 cognitive impairment (0 to 3) (higher values favour control)									

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Usual care	Summary of results	
1 (Bass 2014)	RCT	Serious[2,3,7]	Not serious	N/A	Serious[1]	122	72	MD 0.00 (-0.29, 0.29)	Low
Percentage of participants who had hospital admissions (higher values favour control)									
1 (Bass 2015)	RCT	Serious[2,5]	Not serious	N/A	Serious[9]	206	122	RR 1.27 (0.86, 1.87)	Low
Cognitive symptoms of person living with dementia (higher values favour control)									
2 (Bass 2015, Callahan 2006)	RCT	Serious[2,5]	Not serious	Not serious	Serious[9]	271	171	SMD 0.06 (-0.14, 0.25)	Low
Activities of daily living (of person living with dementia) (higher values favour intervention)									
1 (Callahan 2006)	RCT	Not serious	Not serious	N/A	Serious[1]	65	49	MD 2.30 (-4.48, 9.08)	Moderate
Patient health-related quality of life (higher values favour intervention)									
1 (Vickrey 2006)	RCT	Serious[5,8]	Not serious	N/A	Serious[1]	166	124	MD 0.06 (-0.01, 0.13)	Low
Mean number of physician visits (higher values favour control)									
1 (Bass 2003)	RCT	Serious[2,3,4,]	Not serious	N/A	Serious[1]	92	65	MD 0.01 (-1.35, 1.37)	Low
Behavioural symptoms, such as NPI, of person living with dementia (higher values favour control)									
3 (Bass 2015, Callahan 2006, Chu 2000)	RCT	Serious[2,3,5]	Not serious	Serious[9]	Very serious[10]	304	207	SMD -0.02 (-0.39, 0.36)	Very low
Caregiver relationship strain (Bass 2013) (higher values favour control)									
2 (Bass 2003, Bass 2013)	RCT	Serious[2,3,4]	Not serious	Serious[9]	Very serious[10]	391	252	SMD -0.06 (-0.34, 0.23)	Very low
Caregiver health-related quality of life: mean caregiving attributable health strain (higher values favour intervention)									
1 (Vickrey 2006)	RCT	Serious[5,8]	Not serious	N/A	Serious[1]	166	124	MD 0.01 (-0.04, 0.06)	Low
Caregiver satisfaction with types of services (0 to 3) (higher values favour intervention)									
1 (Bass 2003)	RCT	Serious[2,3,4,]	Not serious	N/A	Serious[1]	92	65	MD 0.02 (-0.18, 0.22)	Low
Caregiver satisfaction with quality of services (different scales used) (higher values favour intervention)									
2 (Bass 2003, Vickrey 2006)	RCT	Serious[2,3,4,5,8]	Not serious	Not serious	Serious[9]	258	189	SMD 0.13 (-0.06, 0.32)	Low
Caregiver satisfaction with information (0 to 3) (higher values favour intervention)									
1 (Bass 2003)	RCT	Serious[2,3,4,]	Not serious	N/A	Serious[9]	92	65	OR 1.15 (0.83, 1.59)	Low

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Usual care	Summary of results	
Caregiver depression (higher values favour control)									
2 (Bass 2003, Fortinsky 2009)	RCT	Serious[2,3,4,5]	Not serious	Not serious	Serious[9]	146	95	SMD -0.23 (-0.49, 0.03)	Low
Caregiver role captivity (0 to 3) (higher values favour control)									
1 (Bass 2003)	RCT	Serious[2,3,4]	Not serious	N/A	Serious[1]	92	65	MD 0.02 (-0.21, 0.25)	Low
Caregiver health-related quality of life (mean EuroQol-5D) (higher values favour intervention)									
1 (Vickrey 2006)	RCT	Serious[5,8]	Not serious	N/A	Serious[1]	166	124	MD 0.01 (-0.04, 0.06)	Low
Behavioural symptoms, such as NPI, of caregiver (higher values favour control)									
1 (Callahan 2006)	RCT	Not serious	Not serious	N/A	Serious[1]	65	49	MD -0.50 (-3.62, 2.62)	Moderate
Caregiver health/symptoms (higher values favour control)									
2 (Bass 2003, Fortinsky 2009)	RCT	Serious[2,3,4,5]	Not serious	Not serious	Very serious[10]	146	95	SMD 0.01 (-0.25, 0.27)	Very low
Caregiver burden (different versions of measurement were used) (higher values favour control)									
2 (Chu 2000, Fortinsky 2009)	RCT	Serious[2,3,5]	Not serious	Serious[9]	Very serious[10]	87	66	SMD -0.19 (-0.73, 0.13)	Very low
Caregiver patient health questionnaire (caregiver's opinion of the health of the person living with dementia) (higher values favour control)									
1 (Callahan 2006)	RCT	Not serious	Not serious	N/A	Serious[1]	65	49	MD -1.50 (-3.34, 0.34)	Moderate
Mean hours of home care services per month (including direct care, case management, respite, personal care assistance and homemaking) from the start of the study to the end of the study (higher values favour control)									
1 (Chu 2000)	RCT	Serious[2,3]	Not serious	N/A	Not serious	33	36	MD 28.60 (0.49, 56.71)	Moderate
Caregiver received as much help as needed with behaviour problem (higher values favour intervention)									
1 (Vickrey 2006)	RCT	Serious[5,8]	Not serious	N/A	Not serious	166	124	MD 15.00 (6.19, 23.81)	Moderate
Symptom management self-efficacy score (how confident the carers are in managing symptoms) (higher values favour intervention)									
1 (Fortinsky 2009)	RCT	Serious[2,3,5]	Not serious	N/A	Serious[1]	54	30	MD -0.34 (-8.92, 8.24)	Low
Support service self-efficacy (how confident are the carers in arranging support services) (higher values favour intervention)									

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Usual care	Summary of results	
1 (Fortinsky 2009)	RCT	Serious[2,3,5]	Not serious	N/A	Serious[1]	54	30	MD 0.70 (-4.13, 5.53)	Low
Caregiver rating of their social support (higher values favour intervention)									
1 (Vickrey 2006)	RCT	Serious[5,8]	Not serious	N/A	Serious[1]	166	124	MD 3.70 (-2.81, 10.27)	Low
Caregiving quality: mean caregiver confidence in caregiving (baseline not measured) (higher values favour intervention)									
1 (Vickrey 2006)	RCT	Serious[5,8]	Not serious	N/A	Not serious	166	124	MD 6.90 (1.94, 11.86)	Moderate
Caregiving quality: mean caregiving mastery (baseline was measured) (higher values favour intervention)									
1 (Vickrey 2006)	RCT	Serious[5,8]	Not serious	N/A	Not serious	166	124	MD 8.70 (2.96, 14.44)	Moderate
Mean number of non-association information and support services (higher values favour control)									
1 (Bass 2003)	RCT	Serious[2,3,4]	Not serious	N/A	Serious[1]	92	65	MD -0.18 (-0.58, 0.22)	Low
Mean number of direct care community services (higher values favour control)									
1 (Bass 2003)	RCT	Serious[2,3,4]	Not serious	N/A	Serious[1]	92	65	MD -0.26 (-0.75, 0.23)	Low
Was there a case management visit during the 1 year period? (0=no, 1=yes) (higher values favour control)									
1 (Bass 2003)	RCT	Serious[2,3,4]	Not serious	N/A	Not serious	92	65	MD -0.16 (-0.29, -0.03)	Moderate

1. Non-significant result
2. The method of randomisation is not given
3. Either no blinding or blinding is not mentioned
4. Baseline data is not provided
5. Not all participants were accounted for
6. $i^2 > 40\%$
7. Not all clinically relevant outcomes were reported
8. It is unclear as to whether the groups were similar at the start of the trial
9. 95% CI crosses one line of a defined MID interval
10. 95% CI crosses two lines of a defined MID interval

Care coordination/management using a protocol/action plan (that involves educating the carers) and approx 10-14 meetings over 4 months vs usual care

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Usual care	Summary of results	
Care recipient Cornell Scale for Depression in Dementia (higher values favour control)									
1 (Lam 2010)	RCT	Not serious	Not serious	N/A	Serious[1]	53	39	MD -0.50 (-3.26, 2.26)	Moderate
Care recipient psychiatric symptoms (NPI) (higher values favour control)									
1 (Lam 2010)	RCT	Not serious	Not serious	N/A	Serious[1]	53	39	MD 5.00 (-10.50, 20.50)	Moderate
Care recipient Personal Well-Being Index-Intellectual Disability (higher values favour intervention)									
1 (Lam 2010)	RCT	Not serious	Not serious	N/A	Serious[1]	53	39	MD 9.30 (-12.27, 30.87)	Moderate
Caregiver Personal Well-Being Index for Adult (higher values favour intervention)									
1 (Lam 2010)	RCT	Not serious	Not serious	N/A	Serious[1]	53	39	MD 2.90 (-9.47, 15.27)	Moderate
Caregiver burden (higher values favour control)									
1 (Lam 2010)	RCT	Not serious	Not serious	N/A	Serious[1]	53	39	MD 1.50 (-14.09, 17.09)	Moderate
Caregiver General Health Questionnaire (mental health assessment) (higher values favour control)									
1 (Lam 2010)	RCT	Not serious	Not serious	N/A	Serious[1]	53	39	MD 1.00 (-3.51, 5.51)	Moderate

1. Non-significant result

Care coordination/management using a protocol/action plan (that involves educating the carers) and 1 meeting per month for 18 months with additional meetings as required vs augmented usual care

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Usual care	Summary of results	
Care recipient total percent unmet care needs (higher values favour control)									
1 (Samus 2014)	RCT	Serious[1]	Not serious	N/A	Not serious	74	114	MD -1.50 (-2.75, -0.25)	Moderate
Person living with dementia's quality of life (QoL-AD) (higher values favour intervention)									
1 (Samus 2014)	RCT	Not serious	Not serious	N/A	Serious[2]	74	114	MD 1.90 (-0.06, 3.86)	Moderate
Person living with dementia's quality of life (ADRQL-40) (higher values favour intervention)									
1 (Samus 2014)	RCT	Not serious	Not serious	N/A	Serious[2]	74	114	MD 0.50 (-2.01, 3.01)	Moderate
Person living with dementia's quality of life (QoL-AD-Informant) (higher values favour intervention)									

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Usual care	Summary of results	
1 (Samus 2014)	RCT	Not serious	Not serious	N/A	Serious[2]	74	114	MD -0.40 (-2.21, 1.41)	Moderate
Care recipient's Cornell Scale for Depression in Dementia (higher values favour control)									
1 (Samus 2014)	RCT	Not serious	Not serious	N/A	Serious[2]	74	114	MD 0.10 (-1.35, 1.55)	Moderate
Care recipient's Neuropsychiatric Inventory – Questionnaire (higher values favour control)									
1 (Samus 2014)	RCT	Not serious	Not serious	N/A	Serious[2]	74	114	MD 0.90 (-0.73, 2.53)	Moderate
Unmet caregiver needs (higher values favour control)									
1 (Tanner 2015)	RCT	Serious[1]	Not serious	N/A	Serious[2]	67	104	MD -0.98 (-4.82, 2.86)	Low
Unmet caregiver education (higher values favour control)									
1 (Tanner 2015)	RCT	Not serious	Not serious	N/A	Serious[2]	67	104	MD -6.98 (-17.56, 3.60)	Moderate
Unmet caregiver resource referral (higher values favour control)									
1 (Tanner 2015)	RCT	Not serious	Not serious	N/A	Serious[2]	67	104	MD -4.45 (-10.91, 2.01)	Moderate
Unmet caregiver mental health care (higher values favour control)									
1 (Tanner 2015)	RCT	Not serious	Not serious	N/A	Serious[2]	67	104	MD -0.39 (-6.98, 6.20)	Moderate
Unmet caregiver medical health care (higher values favour control)									
1 (Tanner 2015)	RCT	Not serious	Not serious	N/A	Serious[2]	67	104	MD 4.51 (-2.01, 11.03)	Moderate
Caregiver QoL: physical health (higher values favour intervention)									
1 (Tanner 2015)	RCT	Not serious	Not serious	N/A	Serious[2]	67	104	MD 1.54 (-1.62, 4.70)	Moderate
Caregiver QoL: mental health (higher values favour intervention)									
1 (Tanner 2015)	RCT	Not serious	Not serious	N/A	Serious[2]	67	104	MD 0.66 (-2.43, 3.75)	Moderate
Caregiver burden (higher values favour control)									
1 (Tanner 2015)	RCT	Not serious	Not serious	N/A	Serious[2]	67	104	MD -1.91 (-4.39, 0.57)	Moderate
Caregiver depression (higher values favour control)									
1 (Tanner 2015)	RCT	Not serious	Not serious	N/A	Serious[2]	67	104	MD -0.39 (-1.25, 0.47)	Moderate
Time spent with care recipient hr/wk ('raw' data) (higher values favour control)									
1 (Tanner 2015)	RCT	Not serious	Not serious	N/A	Not serious	67	104	MD -16.91 (-33.14, -0.68)	High
Caregiver time spent with care recipient hr/wk (after multiple comparison correction) (higher values favour control)									
1 (Tanner 2015)	RCT	Not serious	Not serious	N/A	Serious[2]	67	104	MD 3.16 (-6.74, 13.06)	Moderate

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Usual care	Summary of results	
Caregiver work missed (hours/month) (higher values favour control)									
1 (Tanner 2015)	RCT	Not serious	Not serious	N/A	Serious[2]	67	104	MD -1.41 (-11.79, 8.97)	Moderate
Caregiver difficulty caring for care recipient (higher values favour control)									
1 (Tanner 2015)	RCT	Not serious	Not serious	N/A	Serious[2]	67	104	MD -0.21 (-0.56, 0.14)	Moderate
Overall caregiver health (higher values favour intervention)									
1 (Tanner 2015)	RCT	Not serious	Not serious	N/A	Serious[2]	67	104	MD 0.16 (-0.15, 0.47)	Moderate
Stress from caregiving (higher values favour control)									
1 (Tanner 2015)	RCT	Not serious	Not serious	N/A	Serious2	67	104	MD 0.12 (-0.20, 0.44)	Moderate

1. Not blinded
2. Non-significant result

Care coordination/management using a protocol/action plan (that involves educating the carers) and approx 2 meetings per month for 6 months vs usual care

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Usual care	Summary of results	
Care recipient's MMSE (0 to 30) (higher values favour intervention)									
1 (Chien 2008)	RCT	Not serious	Not serious	N/A	Serious[1]	42	43	MD -0.30 (-2.57, 1.97)	Moderate
Care recipient's Neuro-psychiatric Inventory (different scales were used) (higher values favour control)									
2 (Chien 2008, Dias 2008)	RCT	Not serious	Not serious	Serious[2]	Serious[3]	75	69	SMD -0.95 (-2.07, 0.16)	Moderate
Institutionalisation over the past 6 months - number of times (residential placements or hospitalisations) (higher values favour control)									
1 (Chien 2008)	RCT	Not serious	Not serious	N/A	Not serious	42	43	MD -3.10 (-3.81, -2.39)	High
Institutionalisation over the past 6 months - duration (days per month) (higher values favour control)									
1 (Chien 2008)	RCT	Not serious	Not serious	N/A	Not serious	42	43	MD -6.70 (-8.40, -5.00)	High
Everyday functional abilities of the person living with dementia (higher values favour intervention)									
1 (Dias 2008)	RCT	Not serious	Not serious	N/A	Serious[1]	33	26	MD -0.20 (-1.35, 0.95)	Moderate
Caregiver's 6-item social support questionnaire (0 to 30) (higher values favour intervention)									
1 (Chien 2008)	RCT	Not serious	Not serious	N/A	Not Serious	42	43	MD 1.50 (0.61, 2.39)	High

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Usual care	Summary of results	
Caregiver burden (higher values favour control)									
2 (Chien 2008, Dias 2008)	RCT	Not serious	Not serious	Serious[2]	Serious[3]	75	69	SMD -0.78 (-1.56, -0.00)	Moderate
Caregiver's WHO Quality of Life Scale (28 to 144) (higher values favour intervention)									
1 (Chien 2008)	RCT	Not serious	Not serious	N/A	Not serious	42	43	MD 18.40 (11.48, 25.32)	High
Caregiver mental health (general health questionnaire) (higher values favour control)									
1 (Dias 2008)	RCT	Not serious	Not serious	N/A	Not serious	33	26	MD -2.60 (-4.08, -1.12)	High
Caregiver distress due to problem behaviours (NPIQ-D) (higher values favour control)									
1 (Dias 2008)	RCT	Not serious	Not serious	N/A	Serious[1]	33	26	MD -2.10 (-4.88, 0.68)	Moderate
Family Support Services Index (0 to 16, with higher scores indicating greater varieties of service utilization. We have presented this as a bad thing because of potential cost) (higher values favour control)									
1 (Chien 2008)	RCT	Not serious	Not serious	N/A	Not serious	42	43	MD -1.90 (-2.58, -1.22)	High

1. Non-significant result
2. $i^2 > 40\%$
3. 95% CI crosses one line of a defined MID interval

Care coordination/management using a protocol/action plan (that involves educating the carers) and weekly meetings for a month, followed by a meeting every 2 weeks for 5 months

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Usual care	Summary of results	
MMSE (higher values favour intervention)									
1 (Chien 2011)	RCT	Not serious	Not serious	N/A	Serious[1]	45	45	MD -0.20 (-1.70, 1.30)	Moderate
Neuro-psychiatric Inventory (higher values favour control)									
1 (Chien 2011)	RCT	Not serious	Not serious	N/A	Not serious	45	45	MD -6.80 (-10.89, -2.71)	High
Rate of institutionalisation - number institutionalised during the past 6 months (higher values favour control)									
1 (Chien 2011)	RCT	Not serious	Not serious	N/A	Not serious	45	45	MD -3.00 (-4.00, -2.00)	High
Rate of institutionalisation - duration of institutionalisation (days/month) over the past 6 months (higher values favour control)									
1 (Chien 2011)	RCT	Not serious	Not serious	N/A	Not serious	45	45	MD -4.50 (-7.61, -1.39)	High

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Usual care	Summary of results	
Caregiver WHO Quality of Life (28-144) (higher values favour intervention)									
1 (Chien 2011)	RCT	Not serious	Not serious	N/A	Not serious	45	45	MD 20.50 (15.06, 25.94)	High
Caregiver 6-item social support questionnaire (higher values favour intervention)									
1 (Chien 2011)	RCT	Not serious	Not serious	N/A	Serious[1]	45	45	MD 0.90 (-0.10, 1.90)	Moderate
Family Caregiving Burden Inventory (0-96) (higher values favour control)									
1 (Chien 2011)	RCT	Not serious	Not serious	N/A	Not serious	45	45	MD -19.70 (-24.08, -15.32)	High
Family Support Services Index (responses indicate the number and types of services that families were in need of and receiving) (higher values favour control)									
1 (Chien 2011)	RCT	Not serious	Not serious	N/A	Not serious	45	45	MD -1.50 (-2.16, -0.84)	High

1. Non-significant result

Care coordination by telephone ('experimental') vs care coordination in-person ('control'). Follow-up frequency was monthly for the first 3 months and quarterly thereafter

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Telephone	In-person	Summary of results	
Care-recipient Health Utilities Index (a QoL measure) (higher values favour in-person follow-up)									
1 (Chodosh 2015)	RCT	Serious[1]	Not serious	N/A	Serious[2]	23	20	MD 0.02 (-0.11, 0.15)	Low
Revised Memory and Behaviour Problem Checklist: total number of problems (higher values favour in-person follow-up)									
1 (Chodosh 2015)	RCT	Serious[1]	Not serious	N/A	Serious[2]	23	20	MD 1.07 (-2.28, 4.42)	Low
Caregiver depression (PHQ-9) (higher values favour in-person follow-up)									
1 (Chodosh 2015)	RCT	Serious[1]	Not serious	N/A	Serious[2]	23	20	MD -0.24 (-7.02, 6.54)	Low
Caregiver quality of life: spirituality and faith (higher values favour telephone follow-up)									
1 (Chodosh 2015)	RCT	Serious[1]	Not serious	N/A	Serious[2]	23	20	MD -0.57 (-14.08, 12.94)	Low
Caregiver quality of life: benefits of caregiving (higher values favour in-person follow-up)									

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Telephone	In-person	Summary of results	
1 (Chodosh 2015)	RCT	Serious[1]	Not serious	N/A	Not serious	23	20	MD 5.15 (2.23, 8.07)	Moderate
Caregiver burden (ZBI) (higher values favour in-person follow-up)									
1 (Chodosh 2015)	RCT	Serious[1]	Not serious	N/A	Serious[2]	23	20	MD -0.81 (-10.26, 8.64)	Low

1. By the end of the trial, not all patients were accounted for: 28% of participants became “unreachable” as time progressed
2. Non-significant result

Follow-up organised by memory clinic vs GP

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Usual care	Summary of results	
Patient outcome: QoL-AD, as rated by caregiver (higher values favour memory clinic)									
1 (Meeuwsen 2012)	RCT	Not serious	Not serious	N/A	Serious[1]	78	75	MD 0.49 (-0.65, 1.63)	Moderate
Patient outcome: NPI behaviour (higher values favour GP)									
1 (Meeuwsen 2012)	RCT	Not serious	Not serious	N/A	Serious[1]	78	75	MD 1.13 (-0.51, 2.77)	Moderate
Patient outcome: Interview for Deterioration in Daily living activities in Dementia - help needed (higher values favour GP)									
1 (Meeuwsen 2012)	RCT	Not serious	Not serious	N/A	Serious[1]	78	75	MD 0.66 (-1.88, 3.20)	Moderate
Patient outcome: Interview for Deterioration In Daily living activities in Dementia - take initiative (higher values favour GP)									
1 (Meeuwsen 2012)	RCT	Not serious	Not serious	N/A	Serious[1]	78	75	MD 1.69 (-0.18, 3.56)	Moderate
Patient outcome: Geriatric Depression Scale (higher values favour GP)									
1 (Meeuwsen 2012)	RCT	Not serious	Not serious	N/A	Serious[1]	78	75	MD 0.25 (-0.36, 0.86)	Moderate
Patient outcome: QoL patient (higher values favour memory clinic)									
1 (Meeuwsen 2012)	RCT	Not serious	Not serious	N/A	Serious[1]	78	75	MD 0.25 (-0.74, 1.24)	Moderate

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Usual care	Summary of results	
Caregiver outcome: sense of competence questionnaire (higher values favour memory clinic)									
1 (Meeuwsen 2012)	RCT	Not serious	Not serious	N/A	Serious[1]	78	75	MD -2.43 (-5.82, 0.96)	Moderate
Caregiver outcome: QoL-AD caregiver (higher values favour memory clinic)									
1 (Meeuwsen 2012)	RCT	Not serious	Not serious	N/A	Serious[1]	78	75	MD 0.17 (-0.70, 1.04)	Moderate
Caregiver outcome: Center for Epidemiologic Studies Depression Scale (higher values favour GP)									
1 (Meeuwsen 2012)	RCT	Not serious	Not serious	N/A	Not serious	78	75	MD 2.09 (0.16, 4.02)	High
Caregiver outcome: Inventory for measuring Social Involvement (higher values favour memory clinic)									
1 (Meeuwsen 2012)	RCT	Not serious	Not serious	N/A	Serious[1]	78	75	MD -0.29 (-1.16, 0.58)	Moderate
Caregiver outcome: NPI – emotional (higher values favour GP)									
1 (Meeuwsen 2012)	RCT	Not serious	Not serious	N/A	Serious[1]	78	75	MD 1.43 (-0.94, 3.80)	Moderate
Caregiver outcome: Eysenck Personality Questionnaire (higher values favour GP)									
1 (Meeuwsen 2012)	RCT	Not serious	Not serious	N/A	Serious[1]	78	75	MD 0.68 (0.00, 1.36)	Moderate
Caregiver outcome: State-Trait Anxiety Inventory – trait (higher values favour GP)									
1 (Meeuwsen 2012)	RCT	Not serious	Not serious	N/A	Not serious	78	75	MD 2.14 (0.25, 4.03)	High
Caregiver outcome: State-Trait Anxiety Inventory – state (higher values favour GP)									
1 (Meeuwsen 2012)	RCT	Not serious	Not serious	N/A	Not serious	78	75	MD 2.35 (0.35, 4.35)	High
Caregiver outcome: Pearlin Mastery Scale (higher values favour GP)									
1 (Meeuwsen 2012)	RCT	Not serious	Not serious	N/A	Serious[1]	78	75	MD 0.65 (-0.50, 1.80)	Moderate

1. Non-significant result

The Medicare Alzheimer's Disease Demonstration (care coordination/management with unspecified follow-up frequency) vs usual care

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Usual care	Summary of results	
Hazard ratio for entry into residential care (higher values favour control)									
1 (Miller 1999)	RCT	Serious[1,2,3]	Not serious	N/A	Not serious	4,005	3,798	OR 1.01 (0.92, 1.11)	Moderate
Caregiver burden (higher values favour control)									
1 (Newcomer 1999)	RCT	Serious[5]	Not serious	N/A	Serious[4]	986	920	MD -0.50 (-1.27, 0.27)	Low
Caregiver depression (higher values favour control)									
1 (Newcomer 1999)	RCT	Serious[5]	Not serious	N/A	Serious[4]	986	920	MD -0.32 (-0.64, 0.00)	Low
Likelihood of any caregiver hospitalisation during the study period (a value over 1 favours control)									
1 (Shelton 2001)	RCT	Serious[2,5,6]	Not serious	N/A	Serious[7]	210	202	OR 0.58 (0.35, 0.97)	Low
Likelihood of any caregiver emergency department visit during the study period (a value over 1 favours control)									
1 (Shelton 2001)	RCT	Serious[2,5,6]	Not serious	N/A	Serious[7]	210	202	OR 0.66 (0.40, 1.08)	Low

1. It is unclear as to whether the trial addressed a clearly focused issue because the description of the intervention lacks detail compared to other studies
2. Details of the method of randomisation were not given
3. There is no mention of blinding
4. Non-significant result
5. Not blinded
6. The number of events in either group are not reported. Therefore, only the relative difference is reported, not the absolute difference.
7. 95% CI crosses one line of a defined MID interval

Care coordination/management using DEM-DISC vs care coordination/management without DEM-DISC

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Usual care	Summary of results	
Camberwell Assessment of Needs for the Elderly: total needs (a value over 1 favours control)									
1 (Van Mierlo 2015)	RCT	Serious[1]	Not serious	N/A	Very serious[2]	30	19	OR 0.85 (0.38, 1.31)	Very low
Camberwell Assessment of Needs for the Elderly: total needs (a value under 1 favours control)									

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Usual care	Summary of results	
1 (Van Mierlo 2015)	RCT	Serious[1]	Not serious	N/A	Very serious[2]	30	19	OR 0.81 (0.36, 1.82)	Very low
Camberwell Assessment of Needs for the Elderly: total needs (a value over 1 favours control)									
1 (Van Mierlo 2015)	RCT	Serious[1]	Not serious	N/A	Serious[3]	30	19	OR 1.55 (0.88, 2.75)	Low

1. Blinding is not mentioned, 32% of participants were lost to follow-up, and odds ratios were published so we only know relative differences rather than absolute differences
2. 95% CI crosses two lines of a defined MID interval
3. 95% CI crosses one line of a defined MID interval

Personalised caregiver support for minority groups vs usual care for minority groups

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Usual care	Summary of results	
Caregiver: Short Sense of Competence Questionnaire (higher values favour the intervention)									
1 (Xiao 2016)	RCT	Serious[1]	Not serious	N/A	Not serious	31	30	MD 9.00 (5.78, 12.22)	Moderate
Caregiver: Physical components score (PCS in SF-36) (higher values favour the intervention)									
1 (Xiao 2016)	RCT	Serious[1]	Not serious	N/A	Serious[2]	31	30	MD 2.20 (-1.93, 6.33)	Low
Caregiver: Mental components score (MCS in SF-36) (higher values favour the intervention)									
1 (Xiao 2016)	RCT	Serious[1]	Not serious	N/A	Not serious	31	30	MD 12.70 (8.76, 16.64)	Moderate
Caregiver: Severity of care recipient's BPSD (higher values favour usual care)									
1 (Xiao 2016)	RCT	Serious[1]	Not serious	N/A	Not serious	31	30	MD -3.30 (-6.21, -0.39)	Moderate
Caregiver: Caregiver distress (higher values favour usual care)									
1 (Xiao 2016)	RCT	Serious[1]	Not serious	N/A	Not serious	31	30	MD -6.40 (-11.25, -1.55)	Moderate
Caregiver: Usage of respite care (higher values favour usual care)[3]									
1 (Xiao 2016)	RCT	Serious[1]	Not serious	N/A	Not serious	31	30	MD 1.40 (0.87, 1.93)	Moderate
Caregiver: Satisfaction with service providers (higher values favour the intervention)									
1 (Xiao 2016)	RCT	Serious[1]	Not serious	N/A	Not serious	31	30	MD 22.70 (16.38, 29.02)	Moderate
Caregiver: Usage of community aged care (higher values favour usual care)[3]									
1 (Xiao 2016)	RCT	Serious[1]	Not serious	N/A	Serious[1]	31	30	MD -0.30 (-1.03, 0.43)	Low

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Usual care	Summary of results	

1. Not blinded, randomisation method not given, unclear if both groups were similar at baseline, minority groups differ compared to minority groups in the UK
2. Non-significant result
3. For this review, a greater usage of resources for the effect estimate favours usual care

Care coordination/management using a specific structured protocol vs care coordination/management that is unstructured

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Usual care	Summary of results	
Caregiver's depressive symptoms (higher values favour unstructured coordination)									
1 (Kwak 2011)	RCT	Serious[1]	Not serious	N/A	Serious[2]	41	32	MD 0.15 (-0.14, 0.44)	Low
Caregiver's burden (different scales used) (higher values favour unstructured coordination)									
1 (Kwak 2011)	RCT	Serious[1]	Not serious	N/A	Serious[2]	41	32	MD 0.01 (-0.17, 0.19)	Low
Caregiver identity discrepancy (difference between currently perceived caregiving activities and the caregiver's ideal caregiving activities) (higher values favour unstructured coordination)									
1 (Kwak 2011)	RCT	Serious[1]	Not serious	N/A	Not serious	41	32	MD -0.30 (-0.57, -0.03)	Moderate
Caregiver relationship burden (higher values favour unstructured coordination)									
1 (Kwak 2011)	RCT	Serious[1]	Not serious	N/A	Serious[2]	41	32	MD -0.07 (-0.25, 0.11)	Low
Caregiver stress burden (higher values favour unstructured coordination)									
1 (Kwak 2011)	RCT	Serious[1]	Not serious	N/A	Serious[2]	41	32	MD -0.24 (-0.87, 0.39)	Low

1. Over 70% of care receivers were diagnosed with probable Alzheimer's disease, there was no blinding, and baseline data was not provided so it is not possible to assess whether the two groups were similar at the start.
2. Non-significant result

Case management: combined, by follow-up frequency

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Usual care	Summary of results	
Patient outcome: Cognition, weekly follow-up (higher values favour usual care)									

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Usual care	Summary of results	
1 (Chien 2011)	RCT	Not serious	Not serious	N/A	Very serious[1]	46	46	SMD -0.05 (-0.46, 0.35)	Low
Patient outcome: Cognition, monthly follow-up (higher values favour usual care)									
2 (Bass 2015, Callahan 2006)	RCT	Serious[2,3,4]	Not serious	Not serious	Serious[11]	271	171	SMD 0.06 (-0.14, 0.25)	Low
Patient outcome: Cognition, follow-up every 2 months (higher values favour usual care)									
1 (Chien 2008)	RCT	Not serious	Not serious	N/A	Very serious[1]	42	43	SMD -0.06 (-0.48, 0.37)	Low
Patient outcome: Cognition, all follow-up frequencies (higher values favour usual care)									
4 (Chien 2011, Bass 2015, Callahan 2006, Chien 2008)	RCT	Not serious	Not serious	Not serious	Not serious	359	260	SMD 0.02 (-0.14, 0.18)	High
Depression of the person living with dementia, 10-14 follow-ups over 4 months (higher values favour usual care)									
1 (Lam 2010)	RCT	Not serious	Not serious	N/A	Very serious[1]	53	39	SMD -0.07 (-0.49, 0.34)	Low
Depression of the person living with dementia, monthly follow-ups (higher values favour usual care)									
2 (Callahan 2006, Samus 2014)	RCT	Not serious	Not serious	Not serious	Very serious[1]	139	163	SMD -0.01 (-0.24, 0.22)	Low
Depression of the person living with dementia, all follow-up frequencies (higher values favour usual care)									
3 (Lam 2010, Callahan 2006, Samus 2014)	RCT	Not serious	Not serious	Not serious	Serious[11]	192	202	SMD -0.02 (-0.22, 0.18)	Moderate
QoL of person living with dementia, follow-up every month (which is all follow-up frequencies available) (higher values favour case management)									
2 (Samus 2014, Vickrey 2006)	RCT	Not serious	Not serious	Not serious	Serious[11]	240	238	SMD 0.23 (0.05, 0.42)	Moderate
Behavioural and psychological symptoms of dementia, follow-up every week (higher values favour usual care)									
1 (Chien 2011)	RCT	Not serious	Not serious	N/A	Not serious	46	46	SMD -0.67 (-1.09, -0.25)	High
Behavioural and psychological symptoms of dementia, 10-14 follow-ups over 4 months (higher values favour usual care)									
1 (Lam 2010)	RCT	Not serious	Not serious	N/A	Very serious[1]	53	39	SMD 0.12 (-0.29, 0.54)	Low
Behavioural and psychological symptoms of dementia, monthly follow-ups (higher values favour usual care)									

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Usual care	Summary of results	
4 (Bass 2015, Callahan 2006, Chu 2000, Samus 2014)	RCT	Serious[2,3,5]	Not serious	Serious[6]	Very serious[1]	378	321	SMD 0.03 (-0.25, 0.30)	Very low
Behavioural and psychological symptoms of dementia, follow-ups every 2 months (higher values favour usual care)									
2 (Chien 2008, Dias 2008)	RCT	Not serious	Not serious	Serious[6]	Serious[11]	75	69	SMD -0.95 (-2.07, 0.16)	Low
Behavioural and psychological symptoms of dementia, follow-ups of all frequencies (higher values favour usual care)									
8 (Chien 2011, Lam 2010, Bass 2015, Callahan 2006, Chu 2000, Samus 2014, Chien 2008, Dias 2008)	RCT	Serious[2,3,5]	Not serious	Serious[6]	Serious[11]	552	475	SMD -0.27 (-0.62, 0.09)	Very low
Caregiver depression, follow-ups every month (higher values favour usual care)									
2 (Bass 2003, Tanner 2015)	RCT	Serious[2,7,8]	Not serious	Not serious	Serious[11]	159	169	SMD -0.20 (-0.42, 0.03)	Low
Caregiver depression, unclear frequency of follow-ups (higher values favour usual care)									
1 (Newcomer 1999)	RCT	Serious[2,5,7,9]	Not serious	N/A	Not serious	988	922	SMD -0.09 (-0.18, 0.00)	Moderate
Caregiver depression, all follow-up frequencies (higher values favour usual care)									
3 (Bass 2003, Tanner 2015, Newcomer 1999	RCT	Serious[2,5,7,8,9]	Not serious	Not serious	Not serious	1,147	1,091	SMD -0.10 (-0.19, -0.02)	Moderate
Caregiver burden, follow-ups every week (higher values favour usual care)									
1 (Chien 2011)	RCT	Not serious	Not serious	N/A	Not serious	46	46	SMD -1.82 (-2.31, -1.33)	High
Caregiver burden, 10-14 follow-ups over 4 months (higher values favour usual care)									
1 (Lam 2010)	RCT	Not serious	Not serious	N/A	Very serious[1]	53	39	SMD 0.04 (-0.38, 0.45)	Low
Caregiver burden, follow-ups every month (higher values favour usual care)									

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Usual care	Summary of results	
2 (Chu 2000, Tanner 2015)	RCT	Serious[2,7]	Not serious	Not serious	Serious[11]	100	140	SMD -0.31 (-0.56, -0.05)	Low
Caregiver burden, follow-ups every 2 months (higher values favour usual care)									
2 (Chien 2008, Dias 2008)	RCT	Serious[2,8]	Not serious	Serious[6]	Serious[11]	75	69	SMD -0.78 (-1.56, -0.00)	Very low
Caregiver burden, follow-ups of unclear frequency (higher values favour usual care)									
1 (Newcomer 1999)	RCT	Serious[2,5,7,9]	Not serious	N/A	Not serious	986	920	SMD -0.06 (-0.15, 0.03)	Moderate
Caregiver burden, follow-ups of all frequencies (higher values favour usual care)									
7 (Chien 2011, Lam 2010, Chu 2000, Tanner 2015, Chien 2008, Dias 2008, Newcomer 1999)	RCT	Serious[2,5,7,8,9]	Not serious	Serious[6]	Not serious	1,260	1,214	SMD -0.56 (-0.99, -0.13)	Low
QoL of caregiver, follow-ups every month (higher values favour usual care)									
1 (Vickrey 2006)	RCT	Not serious	Not serious	N/A	Very serious[1]	166	124	SMD 0.02 (-0.21, 0.26)	Low
QoL of caregiver, follow-ups every 2 weeks (higher values favour usual care)									
1 (Chien 2008)	RCT	Not serious	Not serious	N/A	Not serious	42	43	SMD 1.12 (0.66, 1.58)	High
QoL of caregiver, follow-ups every week (higher values favour usual care)									
1 (Chien 2011)	RCT	Not serious	Not serious	N/A	Not serious	46	46	SMD 1.53 (1.06, 2.00)	High
QoL of caregiver, follow-ups of all frequencies (higher values favour usual care)									
3 (Vickrey 2006, Chien 2008, Chien 2011)	RCT	Not serious	Not serious	Serious[6]	Serious[11]	254	213	SMD 0.87 (-0.12, 1.87)	Low
Rate of institutionalisation (number of people institutionalised during the past 6 months), follow-ups every week (higher values favour usual care)									
1 (Chien 2011)	RCT	Not serious	Not serious	N/A	Not serious	46	46	SMD -3.00 (-4.00, -2.00)	High
Rate of institutionalisation (number of people institutionalised during the past 6 months), follow-ups every 2 weeks (higher values favour usual care)									
1 (Chien 2008)	RCT	Not serious	Not serious	N/A	Not serious	42	43	SMD -3.10 (-3.81, -2.39)	High

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Usual care	Summary of results	
Rate of institutionalisation (percentage of people institutionalised – cumulative long-term institutionalisation), follow-ups every 2 months (higher values favour usual care)									
1 (Eloniemi-Sulkava 2009)	RCT	Serious[3,10]	Not serious	N/A	Very serious[1]	63	32	SMD -4.10 (21.69, 13.49)	Very low
Rate of institutionalisation (number of people institutionalised – cumulative long-term institutionalisation), follow-ups of all frequencies (higher values favour usual care)									
3 (Chien 2011, Chien 2008, Eloniemi-Sulkava 2009)	RCT	Serious[3,10]	Not serious	Not serious	Not serious	151	151	SMD -3.07 (-3.65, -2.49)	Moderate

1. 95% CI crosses two lines of a defined MID interval
2. Method of randomisation is not given
3. No blinding
4. Not all clinically significant outcomes were reported
5. High rate of participant attrition
6. $i^2 > 40\%$
7. Blinding is not mentioned
8. Unclear whether both groups were similar at the start of the trail
9. Description of the intervention lacks detail compared to other studies
10. Attrition rates of participants are not mentioned
11. 95% CI crosses one line of a defined MID interval

Case management: combined, by profession of coordinator

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Usual care	Summary of results	
Case management: combined, by profession of coordinator, cognition, mixed professions (higher values favour no case management)									
1 (Bass 2015)	RCT	Serious[1,2,3]	Not serious	N/A	Serious[4]	206	122	SMD 0.08 (-0.14, 0.30)	Low
Case management: combined, by profession of coordinator, cognition, nurse as coordinator (higher values favour no case management)									
3 (Callahan 2006, Chien	RCT	Not serious	Not serious	Not serious	Serious[4]	153	138	SMD -0.04 (-0.27, 0.19)	Moderate

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Usual care	Summary of results	
2008, Chien 2011)									
Case management: combined, by profession of coordinator, cognition, all professions (higher values favour no case management)									
4 (Bass 2015, Callahan 2006, Chien 2008, Chien 2011)	RCT	Serious[1,2,3]	Not serious	Not serious	Not serious	359	260	SMD 0.02 (-0.14, 0.18)	Moderate
Case management: combined, by profession of coordinator, depression of the person living with dementia, nurse (higher values favour no case management)									
1 (Callahan 2006)	RCT	Not serious	Not serious	N/A	Very serious[9]	65	49	SMD -0.05 (-0.42, 0.32)	Low
Case management: combined, by profession of coordinator, depression of the person living with dementia, occupational therapist (higher values favour no case management)									
1 (Lam 2010)	RCT	Not serious	Not serious	N/A	Very serious[9]	53	39	SMD -0.07 (-0.49, 0.34)	Low
Case management: combined, by profession of coordinator, depression of the person living with dementia, social worker (higher values favour no case management)									
1 (Samus 2014)	RCT	Not serious	Not serious	N/A	Very serious[9]	74	114	SMD 0.02 (-0.27, 0.31)	Low
Case management: combined, by profession of coordinator, depression of the person living with dementia, all professions (higher values favour no case management)									
3 (Callahan 2006, Lam 2010, Samus 2014)	RCT	Not serious	Not serious	Not serious	Serious[4]	192	202	SMD -0.02 (-0.22, 0.18)	Moderate
Case management: combined, by profession of coordinator, QoL of person living with dementia, social worker (this is the only group with this outcome) (higher values favour case management)									
2 (Samus 2014, Vickrey 2006)	RCT	Not serious	Not serious	Not serious	Serious[4]	240	238	SMD 0.23 (0.05, 0.42)	Moderate
Case management: combined, by profession of coordinator, behavioural and psychological symptoms of dementia, home care adviser (higher values favour no case management)									
1 (Dias 2008)	RCT	Not serious	Not serious	N/A	Serious[4]	33	26	SMD -0.38 (-0.90, 0.14)	Moderate
Case management: combined, by profession of coordinator, behavioural and psychological symptoms of dementia, mixed professions (higher values favour no case management)									
2 (Bass 2015, Chu 2000)	RCT	Serious[1,2,3,5]	Not serious	Serious[6]	Very serious[9]	239	158	SMD 0.15 (-0.39, 0.70)	Very low
Case management: combined, by profession of coordinator, behavioural and psychological symptoms of dementia, nurse (higher values favour no case management)									

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Usual care	Summary of results	
3 (Callahan 2006, Chien 2008, Chien 2011)	RCT	Not serious	Not serious	Serious[6]	Serious[4]	153	138	SMD -0.83 (-1.49, -0.17)	Low
Case management: combined, by profession of coordinator, behavioural and psychological symptoms of dementia, occupational therapist (higher values favour no case management)									
1 (Lam 2010)	RCT	Not serious	Not serious	N/A	Very serious[9]	53	39	SMD 0.12 (-0.29, 0.54)	Low
Case management: combined, by profession of coordinator, behavioural and psychological symptoms of dementia, social worker (higher values favour no case management)									
1 (Samus 2014)	RCT	Not serious	Not serious	N/A	Serious[4]	74	114	SMD 0.16 (-0.13, 0.45)	Moderate
Case management: combined, by profession of coordinator, behavioural and psychological symptoms of dementia, all professions (higher values favour no case management)									
8 (Dias 2008, Bass 2015, Chu 2000, Callahan 2006, Chien 2008, Chien 2011, Lam 2010, Samus 2014)	RCT	Serious[1,2,3,5]	Not serious	Serious[6]	Serious[4]	552	475	SMD -0.27 (-0.62, 0.09)	Very low
Case management: combined, by profession of coordinator, caregiver depression, nurse (higher values favour no case management)									
1 (Newcomer 1999)	RCT	Serious[1,2,3,7]	Not serious	N/A	Not serious	988	922	SMD -0.09 (-0.18, 0.00)	Moderate
Case management: combined, by profession of coordinator, caregiver depression, social worker (higher values favour no case management)									
2 (Bass 2003, Tanner 2015)	RCT	Not serious	Not serious	Not serious	Serious[4]	159	169	SMD -0.20 (-0.42, 0.03)	Moderate
Case management: combined, by profession of coordinator, caregiver depression, all professions together (higher values favour no case management)									
3 (Newcomer 1999, Bass 2003, Tanner 2015)	RCT	Serious[1,2,3,7]	Not serious	Not serious	Not serious	1,147	1,091	SMD -0.10 (-0.19, -0.02)	Moderate
Case management: combined, by profession of coordinator, caregiver burden, nurse (higher values favour no case management)									

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Usual care	Summary of results	
3 (Chien 2008, Chien 2011, Newcomer 1999)	RCT	Serious[1,2,3,7]	Not serious	Serious[6]	Serious[4]	1,074	1,009	SMD -1.00 (-2.16, 0.16)	Very low
Case management: combined, by profession of coordinator, caregiver burden, occupational therapist (higher values favour no case management)									
1 (Lam 2010)	RCT	Not serious	Not serious	N/A	Very serious[9]	53	39	SMD 0.04 (-0.38, 0.45)	Low
Case management: combined, by profession of coordinator, caregiver burden, mixed (higher values favour no case management)									
1 (Chu 2000)	RCT	Serious[1,5]	Not serious	N/A	Serious[4]	33	36	SMD -0.48 (-0.96, 0.00)	Low
Case management: combined, by profession of coordinator, caregiver burden, home care adviser (higher values favour no case management)									
1 (Dias 2008)	RCT	Not serious	Not serious	N/A	Serious[4]	33	26	SMD -0.37 (-0.89, 0.14)	Moderate
Case management: combined, by profession of coordinator, caregiver burden, social worker (higher values favour no case management)									
1 (Tanner 2015)	RCT	Not serious	Not serious	N/A	Serious[4]	67	104	SMD -0.24 (-0.54, 0.07)	Moderate
Case management: combined, by profession of coordinator, caregiver burden, all professions together (higher values favour no case management)									
7 (Chien 2008, Chien 2011, Newcomer 1999, Lam 2010, Chu 2000, Dias 2008, Tanner 2015)	RCT	Serious[1,2,3,5,7]	Not serious	Serious[6]	Serious[4]	1,260	1,214	SMD -0.56 (-0.99, -0.13)	Very low
Case management: combined, by profession of coordinator, QoL of caregiver, social worker (higher values favour usual care)									
1 (Vickrey 2006)	RCT	Not serious	Not serious	N/A	Very serious[9]	166	124	SMD 0.02 (-0.21, 0.26)	Low
Case management: combined, by profession of coordinator, QoL of caregiver, nurse (higher values favour usual care)									
2 (Chien 2008, Chien 2011)	RCT	Not serious	Not serious	Not serious	Not serious	88	89	SMD 1.32 (0.92, 1.72)	High
Case management: combined, by profession of coordinator, QoL of caregiver, all professions together (higher values favour usual care)									
3 (Vickrey 2006, Chien 2008, Chien 2011)	RCT	Not serious	Not serious	Serious[6]	Serious[4]	254	213	SMD 0.87 (-0.12, 1.87)	Low
Case management: combined, by profession of coordinator, rate of institutionalisation (number of people institutionalised – cumulative long-term institutionalisations or number of institutionalisations over a 6 month period), nurse (which is all professions we have together) (higher values favour usual care)									
3 (Chien 2008, Chien 2011,	RCT	Serious[2,8]	Not serious	Not serious	Not serious	151	151	SMD -3.07 (-3.65, -2.49)	Moderate

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Usual care	Summary of results	
Eloniemi-Sulkava 2009)									

1. Method of randomisation is not given
2. No blinding
3. There was a large attrition rate of participants because of reasons that were not provided
4. 95% CI crosses one line of a defined MID interval
5. Blinding is not mentioned
6. $i^2 > 40\%$
7. The description of the intervention lacks detail compared to other studies
8. Attrition rates of participants are not provided
9. 95% CI crosses two lines of a defined MID interval

Case management: combined, follow-up contact method

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Usual care	Summary of results	
Case management: combined, by follow-up contact method, cognition, clinic follow-up (higher values favour no case management)									
1 (Callahan 2006)	RCT	Not serious	Not serious	N/A	Very serious[1]	65	49	SMD -0.01 (-0.38, 0.36)	Low
Case management: combined, by follow-up contact method, cognition, home visit follow-up (higher values favour no case management)									
2 (Chien 2008, Chien 2011)	RCT	Not serous	Not serious	Not serious	Very serious[1]	88	89	SMD -0.06 (-0.35, 0.24)	Low
Case management: combined, by follow-up contact method, cognition, telephone follow-up (higher values favour no case management)									
1 (Bass 2015)	RCT	Serious[2,3,4]	Not serious	N/A	Serious[10]	206	122	SMD 0.08 (-0.14, 0.30)	Low
Case management: combined, by follow-up contact method, cognition, all follow-up methods combined (higher values favour no case management)									
4 (Callahan 2006, Chien 2008, Chien 2011, Bass 2015)	RCT	Serious[2,3,4]	Not serious	Not serious	Not serious	359	260	SMD 0.02 (-0.14, 0.18)	Moderate
Case management: combined, by follow-up contact method, depression of the person living with dementia, clinic follow-up (higher values favour no case management)									

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Usual care	Summary of results	
1 (Callahan 2006)	RCT	Not serious	Not serious	N/A	Very serious[1]	65	49	SMD -0.05 (-0.42, 0.32)	Low
Case management: combined, by follow-up contact method, depression of the person living with dementia, home visit follow-up (higher values favour no case management)									
1 (Lam 2010)	RCT	Not serious	Not serious	N/A	Very serious[1]	53	39	SMD -0.07 (-0.49, 0.34)	Low
Case management: combined, by follow-up contact method, depression of the person living with dementia, mixed methods follow-up (higher values favour no case management)									
1 (Samas 2014)	RCT	Not serious	Not serious	N/A	Very serious[1]	74	114	SMD 0.02 (-0.27, 0.31)	Low
Case management: combined, by follow-up contact method, depression of the person living with dementia, all follow-up methods results combined (higher values favour no case management)									
3 (Callahan 2006, Lam 2010, Samas 2014)	RCT	Not serious	Not serious	Not serious	Serious[10]	192	202	SMD -0.02 (-0.22, 0.18)	Moderate
Case management: combined, by follow-up contact method, QoL of person living with dementia, mixed follow-up methods (higher values favour case management)									
1 (Samas 2014)	RCT	Not serious	Not serious	N/A	Serious[10]	74	114	SMD 0.29 (-0.01, 0.58)	Moderate
Case management: combined, by follow-up contact method, QoL of person living with dementia, follow-up by telephone (higher values favour case management)									
1 (Vickrey 2006)	RCT	Not serious	Not serious	N/A	Serious[10]	166	124	SMD 0.20 (-0.03, 0.44)	Moderate
Case management: combined, by follow-up contact method, QoL of person living with dementia, all follow-up methods results combined (higher values favour case management)									
2 (Samas 2014, Vickrey 2006)	RCT	Not serious	Not serious	Not serous	Serious[10]	240	238	SMD 0.23 (0.05, 0.42)	Moderate
Case management: combined, by follow-up contact method, behavioural and psychological symptoms of dementia, clinic follow-up (higher values favour no case management)									
2 (Callahan 2006, Dias 2008)	RCT	Not serious	Not serious	Not serious	Serious[10]	98	75	SMD -0.35 (-0.65, -0.05)	Moderate
Case management: combined, by follow-up contact method, behavioural and psychological symptoms of dementia, home visit follow-up (higher values favour no case management)									
4 (Chien 2008, Chien 2011, Chu 2000, Lam 2010)	RCT	Serious[2,5]	Not serious	Serious[6]	Very serious[1]	174	164	SMD -0.40 (-1.22, 0.43)	Very low

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Usual care	Summary of results	
Case management: combined, by follow-up contact method, behavioural and psychological symptoms of dementia, mixed methods follow-up (higher values favour no case management)									
1 (Samas 2014)	RCT	Not serious	Not serious	N/A	Serious[10]	74	114	SMD 0.16 (-0.13, 0.45)	Moderate
Case management: combined, by follow-up contact method, behavioural and psychological symptoms of dementia, telephone follow-up (higher values favour no case management)									
1 (Bass 2015)	RCT	Serious[2,3,4]	Not serious	N/A	Serious[10]	206	122	SMD -0.09 (-0.31, 0.14)	Low
Case management: combined, by follow-up contact method, behavioural and psychological symptoms of dementia, all follow-up methods results combined (higher values favour no case management)									
8 (Callahan 2006, Dias 2008, Chien 2008, Chien 2011, Chu 2000, Lam 2010, Samas 2014, Bass 2015)	RCT	Serious[2,3,4,5]	Not serious	Serious[6]	Serious[10]	552	475	SMD -0.27 (-0.62, 0.09)	Very low
Case management: combined, by follow-up contact method, caregiver depression, home visit follow-up (higher values favour no case management)									
1 (Newcomer 1999)	RCT	Serious[2,4,5,7]	Not serious	N/A	Not serious	988	922	SMD -0.09 (-0.18, 0.00)	Moderate
Case management: combined, by follow-up contact method, caregiver depression, mixed follow-up methods (higher values favour no case management)									
1 (Tanner 2015)	RCT	Not serious	Not serious	N/A	Serious[10]	67	104	SMD -0.14 (-0.44, 0.17)	Moderate
Case management: combined, by follow-up contact method, caregiver depression, telephone follow-up (higher values favour no case management)									
1 (Bass 2003)	RCT	Serious[2,5,8]	Not serious	N/A	Serious[10]	92	65	SMD -0.26 (-0.58, 0.06)	Low
Case management: combined, by follow-up contact method, caregiver depression, all follow-up methods results combined (higher values favour no case management)									
3 (Newcomer 1999, Tanner 2015, Bass 2003)	RCT	Serious[2,5,8]	Not serious	Not serious	Not serious	1147	1091	SMD -0.10 (-0.19, -0.02)	Moderate
Case management: combined, by follow-up contact method, caregiver burden, clinic follow-up (higher values favour no case management)									
1 (Dias 2008)	RCT	Not serious	Not serious	N/A	Serious[10]	33	26	SMD -0.37 (-0.89, 0.14)	Moderate
Case management: combined, by follow-up contact method, caregiver burden, home visit follow-up (higher values favour no case management)									

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Usual care	Summary of results	
4 (Chien 2008, Chien 2011, Chu 2000, Lam 2010)	RCT	Serious[2,5]	Not serious	Serious[6]	Serious[10]	1,160	1,084	SMD -0.68 (-1.32, -0.04)	Very low
Case management: combined, by follow-up contact method, caregiver burden, mixed follow-up (higher values favour no case management)									
1 (Tanner 2015)	RCT	Not serious	Not serious	N/A	Serious[10]	67	104	SMD -0.24 (-0.54, 0.07)	Moderate
Case management: combined, by follow-up contact method, caregiver burden, all follow-up methods results combined (higher values favour no case management)									
6 (Dias 2008, Chien 2008, Chien 2011, Chu 2000, Lam 2010, Tanner 2015)	RCT	Serious[2,5]	Not serious	Serious[6]	Serious[10]	1,260	1,214	SMD -0.56 (-0.99, -0.13)	Very low
Case management: combined, by follow-up contact method, QoL of caregiver, home visit follow-up (higher values favour no case management)									
2 (Chien 2008, Chien 2011)	RCT	Not serous	Not serious	Not serious	Not serious	88	89	SMD 1.32 (0.92, 1.72)	High
Case management: combined, by follow-up contact method, QoL of caregiver, telephone follow-up (higher values favour no case management)									
1 (Vickrey 2006)	RCT	Not serious	Not serious	N/A	Very serious[1]	166	124	SMD 0.02 (-0.21, 0.26)	Low
Case management: combined, by follow-up contact method, QoL of caregiver, all follow-up methods results combined (higher values favour no case management)									
3 (Chien 2008, Chien 2011, Vickrey 2006)	RCT	Not serious	Not serious	Serious[6]	Serious[10]	254	213	SMD 0.87 (-0.12, 1.87)	Low
Case management: combined, by follow-up contact method, rate of institutionalisation (number of people institutionalised over a 6-month period), home visit follow-up (higher values favour no case management)									
2 (Chien 2008, Chien 2011)	RCT	Not serous	Not serious	Not serious	Not serious	88	89	SMD -3.07 (-3.65, -2.49)	High
Case management: combined, by follow-up contact method, rate of institutionalisation (number of people institutionalised – cumulative long-term institutionalisations), mixed follow-up (higher values favour no case management)									
1 (Eloniemi-Sulkava 2009)	RCT	Serious[3,9]	Not serious	N/A	Very serious[1]	63	62	SMD -4.10 (-21.69, 13.49)	Very low
Case management: combined, by follow-up contact method, rate of institutionalisation (number of people institutionalised, cumulative long-term institutionalisations or number of institutionalisations over a 6-month period), all follow-up methods results combined (higher values favour no case management)									

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Usual care	Summary of results	
3 (Chien 2008, Chien 2011, Eloniemi-Sulkava 2009)	RCT	Serious[3,9]	Not serious	Not serious	Not serious	151	151	SMD -3.07 (-3.65, -2.49)	Moderate

1. 95% CI crosses two lines of a defined MID interval
2. Method of randomisation is not given
3. No blinding
4. Large rate of participant attrition with no explanation
5. Blinding not mentioned
6. $i^2 > 40\%$
7. The description of the intervention lacks detail compared to other studies
8. Unclear whether both groups were similar at the start of the trail because baseline data is not provided
9. Attrition rates of participants are not mentioned
10. 95% CI crosses one line of a defined MID interval

Case management: combined, by country

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Usual care	Summary of results	
Case management: combined, by country, cognition, Hong Kong (higher values favour no case management)									
2 (Chien 2008, Chien 2011)	RCT	Not serous	Not serious	Not serious	Very serious[1]	88	89	SMD -0.06 (-0.35, 0.24)	Low
Case management: combined, by country, cognition, USA (higher values favour no case management)									
2 (Bass 2015, Callahan 2006)	RCT	Serious[2,3,4]	Not serious	N/A	Serious[10]	271	171	SMD 0.06 (-0.14, 0.25)	Low
Case management: combined, by country, cognition, all follow-up methods results combined (higher values favour no case management)									
4 (Chien 2008, Chien 2011, Bass 2015, Callahan 2006)	RCT	Serious[2,3,4]	Not serious	Not serious	Not serious	359	260	SMD 0.02 (-0.14, 0.18)	Moderate
Case management: combined, by country, depression of the person living with dementia, Hong Kong (higher values favour no case management)									

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Usual care	Summary of results	
1 (Lam 2010)	RCT	Not serious	Not serious	N/A	Very serious[1]	53	39	SMD -0.07 (-0.49, 0.34)	Low
Case management: combined, by country, depression of the person living with dementia, USA (higher values favour no case management)									
2 (Callahan 2006, Samus 2014)	RCT	Not serious	Not serious	Not serious	Very serious[1]	139	163	SMD -0.01 (-0.24, 0.22)	Low
Case management: combined, by country, depression of the person living with dementia, all follow-up methods results combined (higher values favour no case management)									
3 (Lam 2010, Callahan 2006, Samus 2014)	RCT	Not serious	Not serious	Serious[6]	Serious[10]	192	202	SMD -0.02 (-0.22, 0.18)	Low
Case management: combined, by country, QoL of the person living with dementia, USA (which is all follow-up methods results combined) (higher values favour no case management)									
2 (Samus 2014, Vickrey 2006)	RCT	Not serious	Not serious	Not serious	Serious[10]	240	238	SMD 0.23 (0.05, 0.42)	Moderate
Case management: combined, by country, behavioural and psychological symptoms of dementia, Canada (higher values favour no case management)									
1 (Chu 2000)	RCT	Serious[2,6]	Not serious	N/A	Serious[10]	33	36	SMD 0.48 (-0.00, 0.96)	Low
Case management: combined, by country, behavioural and psychological symptoms of dementia, Hong Kong (higher values favour no case management)									
3 (Chien 2008, Chien 2011, Lam 2010)	RCT	Not serious	Not serious	Serious[6]	Very serious[1]	141	128	SMD -0.68 (-1.59, 0.22)	Very low
Case management: combined, by country, behavioural and psychological symptoms of dementia, India (higher values favour no case management)									
1 (Dias 2008)	RCT	Not serious	Not serious	N/A	Serious[10]	33	26	SMD -0.38 (-0.90, 0.14)	Moderate
Case management: combined, by country, behavioural and psychological symptoms of dementia, USA (higher values favour no case management)									
3 (Bass 2015, Callahan 2006, Samus 2014)	RCT	Serious[2,3,4]	Not serious	Serious[6]	Serious[10]	345	285	SMD -0.07 (-0.32, 0.18)	Very low
Case management: combined, by country, behavioural and psychological symptoms of dementia, all countries combined (higher values favour no case management)									
8 (Chu 2000, Chien 2008, Chien 2011, Lam 2010, Dias 2008,	RCT	Serious[2,3,4]	Not serious	Serious[6]	Serious[10]	552	475	SMD -0.27 (-0.62, 0.09)	Very low

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Usual care	Summary of results	
Bass 2015, Callahan 2006, Samus 2014)									
Case management: combined, by country, caregiver depression, USA (which is all countries combined) (higher values favour no case management)									
3 (Bass 2003, Newcomer 1999, Tanner 2015)	RCT	Serious[2,4,7]	Not serious	Not serious	Not serious	1,147	1,091	SMD -0.10 (-0.19, -0.02)	Moderate
Case management: combined, by country, caregiver burden, Canada (higher values favour no case management)									
1 (Chu 2000)	RCT	Serious[2,6]	Not serious	N/A	Serious[10]	33	36	SMD -0.48 (-0.96, 0.00)	Low
Case management: combined, by country, caregiver burden, Hong Kong (higher values favour no case management)									
3 (Chien 2008, Chien 2011, Lam 2010)	RCT	Not serious	Not serious	Serious[6]	Serious[10]	141	128	SMD -0.98 (-2.07, 0.11)	Low
Case management: combined, by country, caregiver burden, India (higher values favour no case management)									
1 (Dias 2008)	RCT	Not serious	Not serious	N/A	Serious[10]	33	26	SMD -0.37 (-0.89, 0.14)	Moderate
Case management: combined, by country, caregiver burden, USA (higher values favour no case management)									
2 (Newcomer 1999, Tanner 2015)	RCT	Serious[2,6,8]	Not serious	Not serious	Serious[10]	1053	1024	SMD -0.08 (-0.20, 0.04)	Low
Case management: combined, by country, caregiver burden, all countries combined (higher values favour no case management)									
7 (Chu 2000, Chien 2008, Chien 2011, Lam 2010, Dias 2008, Newcomer 1999, Tanner 2015)	RCT	Serious[2,6,8]	Not serious	Serious[6]	Serious[10]	1,260	1,214	SMD -0.56 (-0.99, -0.13)	Very low
Case management: combined, by country, QoL of caregiver, Hong Kong (higher values favour no case management)									
2 (Chien 2008, Chien 2011)	RCT	Not serous	Not serious	Not serious	Not serious	88	89	SMD 1.32 (0.92, 1.72)	High
Case management: combined, by country, QoL of caregiver, USA (higher values favour no case management)									
1 (Vickrey 2006)	RCT	Not serious	Not serious	N/A	Very serious[1]	166	124	SMD 0.02 (-0.21, 0.26)	Low

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Usual care	Summary of results	
Case management: combined, by country, QoL of caregiver, all countries combined (higher values favour no case management)									
3 (Chien 2008, Chien 2011, Vickrey 2006)	RCT	Not serious	Not serious	Serious[6]	Serious[10]	254	213	SMD 0.87 (-0.12, 1.87)	Low
Case management: combined, by country, rate of institutionalisation (number of people institutionalised – cumulative long-term institutionalisations), Finland (higher values favour no case management)									
1 (Eloniemi-Sulkava 2009)	RCT	Serious[3,9]	Not serious	N/A	Very serious[1]	63	62	SMD -4.10 (-21.69, 13.49)	Very low
Case management: combined, by country, rate of institutionalisation (number of people institutionalised – number of institutionalisations over a 6-month period), Hong Kong (higher values favour no case management)									
2 (Chien 2008, Chien 2011)	RCT	Not serous	Not serious	Not serious	Not serious	88	89	SMD -3.07 (-3.65, -2.49)	High
Case management: combined, by country, rate of institutionalisation (number of people institutionalised – cumulative long-term institutionalisations and number of institutionalisations over a 6-month period), all countries combined (higher values favour no case management)									
3 (Eloniemi-Sulkava 2009, Chien 2008, Chien 2011)	RCT	Serious[3,9]	Not serious	Not serious	Not serious	151	151	SMD -3.07 (-3.65, -2.49)	Moderate

1. 95% CI crosses two lines of a defined MID interval
2. Method of randomisation is not given
3. No blinding
4. Large rate of participant attrition with no explanation
5. $i^2 > 40\%$
6. Blinding is not mentioned
7. Unclear whether both groups were similar at the start of the trail because baseline data is not provided
8. The description of the intervention lacks detail compared to other studies
9. Attrition rates of participants are not mentioned
10. 95% CI crosses one line of a defined MID interval

G.3.2 Post diagnosis review for people living with dementia

- How should people living with dementia be reviewed post diagnosis?

G.3.2.1 Managed health services in partnership with Alzheimer's associations services versus usual managed care services only

Quality assessment						No of patients	Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision			
Outcome: Number of emergency department visits at 12 months								
Bass (2003)	RCT	Not serious	Not serious	N/A	Serious[1]	157	MD -0.17 (-0.51, 0.17)	Moderate
Outcome: Number of hospital admissions at 12 months								
Bass (2003)	RCT	Not serious	Not serious	N/A	Serious[1]	157	MD -0.08 (-0.26, 0.10)	Moderate
Outcome: Number of physician visits at 12 months								
Bass (2003)	RCT	Not serious	Not serious	N/A	Serious[1]	157	MD 0.01 (-1.36, 1.38)	Moderate
Outcome: Use of case management at 12 months								
Bass (2003)	RCT	Not serious	Not serious	N/A	Not serious	157	MD -0.16 (-0.29, -0.03)	High
Outcome: Use of direct care community services at 12 months								
Bass (2003)	RCT	Not serious	Not serious	N/A	Serious[1]	157	MD 0.02 (-0.47, 0.51)	Moderate
Outcome: Use of non-Association information and support services								
Bass (2003)	RCT	Not serious	Not serious	N/A	Serious[1]	157	MD -0.10 (-0.50, 0.30)	Moderate

1. Non-significant result

G.3.2.2 Multidisciplinary case conferences versus usual care

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention Medication advisory case conference	Comparator Usual care		
Outcome: Medicines Appropriation Index at 3 months									

Quality assessment						No of patients			
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention Medication advisory case conference	Comparator Usual care	Effect estimate	Quality
Crotty (2004)	RCT	Not serious	Serious[1]	N/A	Serious[2]	50	54	MD 0.20 (-2.74, 3.14)	Low
Outcome: Change in Medicines Appropriation Index scores at 3 months									
Crotty (2004)	RCT	Not serious	Serious[1]	N/A	Not serious	50	54	MD 3.69 (1.53, 5.85)	Moderate
Outcome: Number of drugs at 3 months									
Crotty (2004)	RCT	Not serious	Serious[1]	N/A	Serious[2]	50	54	MD -0.20 (-1.56, 1.16)	Low
Outcome: Change in number of drugs at 3 months									
Crotty (2004)	RCT	Not serious	Serious[1]	N/A	Serious[2]	50	54	MD 0.39 (-0.55, 1.33)	Low
Outcome: Nursing Home Behaviour Problem Checklist at 3 months									
Crotty (2004)	RCT	Not serious	Serious[1]	N/A	Serious[2]	50	54	MD -10.90 (-27.87, 6.07)	Low
Outcome: Change in Nursing Home Behaviour Problem Checklist at 3 months									
Crotty (2004)	RCT	Not serious	Serious[1]	N/A	Serious[2]	50	54	MD -2.70 (-14.97, 9.57)	Low

1. Population were aged care residents with problem-behaviours and medication problems (including people living with dementia)
2. Non-significant result

Quality assessment						No of patients			
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention Within facility control [a]	Comparator Control group [a]	Effect estimate	Quality
Outcome: Medicines Appropriation Index at 3 months									
Crotty (2004)	RCT	Not serious	Serious[1]	N/A	Serious[2]	50	54	MD 2.50 (-0.47, 5.47)	Low
Outcome: Change in Medicines Appropriation Index scores at 3 months									
Crotty (2004)	RCT	Not serious	Serious[1]	N/A	Serious[2]	50	54	MD -0.53 (-2.06, 1.00)	Low
Outcome: Number of drugs at 3 months									
Crotty (2004)	RCT	Not serious	Serious[1]	N/A	Serious[2]	50	54	MD 0.40 (-0.77, 1.57)	Low
Outcome: Change in number of drugs at 3 months									
Crotty (2004)	RCT	Not serious	Serious[1]	N/A	Serious[2]	50	54	MD -0.24([-1.06, 0.58)	Low
Outcome: Nursing Home Behaviour Problem Checklist at 3 months									
Crotty (2004)	RCT	Not serious	Serious[1]	N/A	Serious[2]	50	54	MD -12.90 (-28.92, 3.12)	Low
Outcome: Change in Nursing Home Behaviour Problem Checklist at 3 months									
Crotty (2004)	RCT	Not serious	Serious[1]	N/A	Serious[2]	50	54	MD -3.00 (-10.52, 4.52)	Low

1. Population were aged care residents with problem-behaviours and medication problems (including people living with dementia)
2. Non-significant result

a. Comparison to reflect any carry-over effect for residents not discussed in case conferences

G.3.2.3 Network multidisciplinary care versus usual care

Quality assessment						No of patients			
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention (multidisciplinary care)	Comparator (usual care)	Effect estimate	Quality
Outcome: Functional outcomes (NAA) at 12 months (lower values=better functional ability)									
Kohler (2014)	RCT	Serious[1]	Not serious	N/A	Serious[2]	97	106	MD 0.50 (-1.68, 2.68)	Low
Outcome: Functional outcomes IADL at 12 months (higher values= better functioning)									
Kohler (2014)	RCT	Serious[1]	Not serious	N/A	Serious[2]	97	106	MD -0.10 (-0.66 0.46)	Low
Outcome: Cognition MMSE (higher values= better cognitive functioning)									
Kohler (2014)	RCT	Serious[1]	Not serious	N/A	Serious[2]	97	106	MD 0.50 (-1.23, 2.23)	Low
Outcome: Health related quality of life (EQ5D VAS) at 12 months (higher values= better rating)									
Kohler (2014)	RCT	Serious[1]	Not serious	N/A	Serious[2]	97	106	MD -1.10 (-6.64, 4.44)	Low
Outcome: Quality of life (QoL-AD) at 12 months (higher values= better quality of life)									
Kohler (2014)	RCT	Serious[1]	Not serious	N/A	Serious[2]	97	106	MD 0.20 (-1.36, 1.76)	Low
Outcome: Caregiver Health related quality of life (EQ5D VAS) at 12 months (higher values= better rating)									
Kohler (2014)	RCT	Serious[1]	Not serious	N/A	Serious[2]	97	106	MD 0.50 (-4.70, 5.70)	Low
Outcome: SF-36 Health survey Physical health sum score at 12 months (higher values = better rating)									
Kohler (2014)	RCT	Serious[1]	Not serious	N/A	Serious[2]	97	106	MD 2.60 (-0.81, 6.01)	Low
Outcome: SF-36 Health survey Mental health sum score at 12 months (higher values = better rating)									
Kohler (2014)	RCT	Serious[1]	Not serious	N/A	Serious[2]	97	106	MD 0.10 (-2.67, 2.87)	Low

1. High risk of bias due to un-blinded allocation and assignment to intervention groups

Quality assessment						No of patients			
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention (multidisciplinary care)	Comparator (usual care)	Effect estimate	Quality
2. Non-significant result									

G.3.2.4 Memory clinic follow up versus GP follow up

Quality assessment						No of patients	Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision			
Outcome: QoL-AD (patient, as reported by caregiver) at 12 months (higher values favours intervention)								
Meeuwsen (2012)	RCT	Not serious	Not serious	N/A	Serious[1]	153	MD 0.49 (-0.66, 1.63)	Moderate
Outcome: QoL-AD (patient report) at 12 months (higher values= favours intervention)								
Meeuwsen (2012)	RCT	Not serious	Not serious	N/A	Serious[1]	145	MD 0.25 (-0.76, 1.23)	Moderate
Outcome: NPI behaviour at 12 months (lower values favours intervention)								
Meeuwsen (2012)	RCT	Not serious	Not serious	N/A	Serious[1]	152	MD 1.13 (-0.51, 2.77)	Moderate
Outcome: Interview for deterioration in daily living activities in dementia (help needed) at 12 months								
Meeuwsen (2012)	RCT	Not serious	Not serious	N/A	Serious[1]	148	MD 0.66 (-1.88, 3.20)	Moderate
Outcome: Interview for deterioration in daily living in dementia (take initiative) at 12 months								
Meeuwsen (2012)	RCT	Not serious	Not serious	N/A	Serious[1]	152	MD 1.69 (-0.18, 3.56)	Moderate
Outcome: Geriatric Depression Scale at 12 months								
Meeuwsen (2012)	RCT	Not serious	Not serious	N/A	Serious[1]	139	MD 0.25 (-0.36, 0.86)	Moderate
Outcome: Caregivers Sense of Competence at 12 months								
Meeuwsen (2012)	RCT	Not serious	Not serious	N/A	Serious[1]	153	MD -2.43 (-5.82, 0.96)	Moderate
Outcome: Caregivers QoL-AD at 12 months								
Meeuwsen (2012)	RCT	Not serious	Not serious	N/A	Serious[1]	154	MD 0.17 (-0.70, 1.04)	Moderate
Outcome: Caregivers CES Depression at 12 months								
Meeuwsen (2012)	RCT	Not serious	Not serious	N/A	Not serious	151	MD 2.09 (0.15, 4.02)	High
Outcome: Caregivers Inventory for measuring social involvement at 12 months								

Quality assessment						No of patients	Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision			
Meeuwsen (2012)	RCT	Not serious	Not serious	N/A	Serious[1]	151	MD -0.29 (-0.97, 0.78)	Moderate
Outcome: Caregivers NPI (emotional) at 12 months								
Meeuwsen (2012)	RCT	Not serious	Not serious	N/A	Serious[1]	152	MD 1.43 (-0.94, 3.80)	Moderate
Outcome: Caregivers Eysenck personality questionnaire at 12 months								
Meeuwsen (2012)	RCT	Not serious	Not serious	N/A	Serious[1]	151	MD 0.68 (-0.01, 1.36)	Moderate
Outcome: Caregivers State trait anxiety inventory (trait) at 12 months								
Meeuwsen (2012)	RCT	Not serious	Not serious	N/A	Not serious	152	MD 2.14 (0.24, 4.03)	High
Outcome: Caregivers State trait anxiety inventory (state) at 12 months								
Meeuwsen (2012)	RCT	Not serious	Not serious	N/A	Not serious	151	MD 2.35 (0.35, 4.36)	High
Outcome: Caregivers Pearlin Mastery scale at 12 months								
Meeuwsen (2012)	RCT	Not serious	Not serious	N/A	Serious[1]	152	MD 0.65 (-0.51, 1.80)	Moderate

1. Non-significant result

G.3.2.5 Specialist care in memory clinic versus usual care in memory clinic

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention (specialist care in memory clinic)	Comparator (usual care in memory clinic)		
Outcome: Functional decline at 2 years (ADCS-ADL)									
Nourhashemi (2010)	RCT	Serious[1]	Not serious	N/A	Serious[2]	224	257	MD 1.00 (-2.27, 4.27)	Low
Outcome: Mean time to admission at 2 years (mean number of days)									
Nourhashemi (2010)	RCT	Serious[1]	Not serious	N/A	Serious[2]	224	257	MD 3.10 (-33.27, 39.47)	Low

Quality assessment						No of patients			
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention (specialist care in memory clinic)	Comparator (usual care in memory clinic)	Effect estimate	Quality
Outcome: Risk of admission to care									
Nourhashemi (2010)	RCT	Serious[1]	Not serious	N/A	Very serious[3]	224	257	HR 0.95 (0.67, 1.36)	Very low
Outcome: Risk of mortality									
Nourhashemi (2010)	RCT	Serious[1]	Not serious	N/A	Very serious[3]	224	257	HR 0.80 (0.51, 1.25)	Very low
Outcome: Admissions due to worsening conditions									
Nourhashemi (2010)	RCT	Serious[1]	Not serious	N/A	Not serious	224	257	RR 0.62 (0.52, 0.76)	Moderate
Outcome: Admissions due to caregiver reasons									
Nourhashemi (2010)	RCT	Serious[1]	Not serious	N/A	Not serious	181/257 (70.59%)	66/224 (29.41%)	RR 2.39 (1.92, 2.97)	Moderate

1. Large numbers of loss to follow up at 2 years
2. Non-significant result
3. 95% CI crosses two lines of a defined MID interval

G.4 Inpatient care

G.4.1 Caring for people living with dementia who are admitted to hospital

- How should people living with dementia be cared for when admitted to hospital?

G.4.1.1 Nurse-led mental health liaison service versus usual care

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Comparator		
Outcome: Geriatric Depression Scale (follow up at 6-8 weeks									
Baldwin (2004)	RCT	Not serious	Serious[1]	N/A	Serious[2]	54	60	MD -1.80 (-4.15, 0.55)	Low
Outcome: MMSE at 6-8 weeks									
Baldwin (2004)	RCT	Not serious	Serious[1]	N/A	Serious[2]	57	61	MD -1.50 (-4.02, 1.02)	Low
Outcome: Length of stay in hospital (days)									
Baldwin (2004)	RCT	Not serious	Serious[1]	N/A	Serious[2]	77	76	MD -1.70 (-11.00, 7.60)	Low
Outcome: Health of Nation Outcome scale (65+ scores)									
Baldwin (2004)	RCT	Not serious	Serious[1]	N/A	Serious[2]	58	59	MD 0.00 (-1.75; 1.75)	Low
Outcome: Prescribed psychotropic medicine at discharge									
Baldwin (2004)	RCT	Not serious	Serious[1]	N/A	Very serious[3]	26/59 (44%)	27/64 (42%)	RR 1.04 (0.70, 1.57)	Very low
Outcome: Readmissions at 3 months									
Baldwin (2004)	RCT	Not serious	Serious[1]	N/A	Very serious[3]	19/77 (24.7%)	21/76 (27.6%)	RR 0.89 (0.52, 1.52)	Very low
Outcome: Deaths at 3 months									

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Desig n	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Comparator		
Baldwin (2004)	RCT	Not serious	Serious[1]	N/A	Serious[2]	17/77 (22.1%)	13/76 (17.1%)	RR 1.29 (0.68, 2.47)	Low

1. Mixed population of people with depression and cognitive impairment at baseline.
2. Non-significant result.
3. 95% CI crosses two lines of a defined MID interval

G.4.1.2 Family-centred function focused care versus usual care

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Interventio n	Comparator		
Outcome: Mean difference in length of stay at discharge									
Boltz (2015)	Non randomised controlled trial	Very serious[1]	Not serious	N/A	Serious[2]	44	42	MD -0.40 (-1.27, 0.47)	Very low
Outcome: Hospital readmissions at 30 days									
Boltz (2015)	Non randomised controlled trial	Very serious[1]	Not serious	N/A	Serious[2]	44	42	MD -7.00 (-14.55, 0.55)	Very low
Outcome: Utilisation of post-acute rehabilitation at discharge									
Boltz (2015)	Non randomised controlled trial	Very serious[1]	Not serious	N/A	Serious[2]	44	42	MD 2.00 (-25.48, 29.48)	Very low
Outcome: Activities of Daily Living (Barthel Index) at 2 months									
Boltz (2015)	Non randomised controlled trial	Very serious[1]	Not serious	N/A	Not serious	44	42	MD 20.7 (10.32, 31.08)	Low
Outcome: Walking performance (50 yards) at 2 months									
Boltz (2015)	Non randomised controlled trial	Very serious[1]	Not serious	N/A	Not serious	44	42	MD 5.60 (3.39, 7.81)	Low
Outcome: Gait and Balance (Tinetti Scale) at 2 months									

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Comparator		
Boltz (2015)	Non randomised controlled trial	Very serious[1]	Not serious	N/A	Serious[2]	44	42	MD 1.50 (-2.39, 5.39)	Very low
Outcome: Delirium severity (Delirium severity Scale) at 2 months									
Boltz (2015)	Non randomised controlled trial	Very serious[1]	Not serious	N/A	Not serious	44	42	MD -2.00 (-3.09, -0.91)	Low
Outcome: Delirium present at 2 months post discharge									
Boltz (2015)	Non randomised controlled trial	Very serious[1]	Not serious	N/A	Not serious	44	42	MD -9.00 (-17.83, -0.17)	Low
Outcome: Carer preparedness for caregiving at 2 months									
Boltz (2015)	Non randomised controlled trial	Very serious[1]	Not serious	N/A	Serious[2]	44	42	MD -3.10 (-5.73, 0.47)	Very low
Outcome: Carer anxiety (HADS-A) at 2 months									
Boltz (2015)	Non randomised controlled trial	Very serious[1]	Not serious	N/A	Serious[2]	44	42	MD -1.60 (-3.57, 0.37)	Very low
Outcome: Carer depression (HADS-D) at 2 months									
Boltz (2015)	Non randomised controlled trial	Very serious[1]	Not serious	N/A	Serious[2]	44	42	MD -0.70 (-2.54, 1.14)	Very low
Outcome: Carer role strain (Modified Caregiver Strain Index) at 2 months									
Boltz (2015)	Non randomised controlled trial	Very serious[1]	Not serious	N/A	Serious[2]	44	42	MD -0.80 (-3.06, 1.46)	Very low
Outcome: Carer mutuality at 2 months									
Boltz (2015)	Non randomised controlled trial	Very serious[1]	Not serious	N/A	Serious[2]	44	42	MD 3.50 (-1.51, 8.51)	Very low

1. Non-randomised study; high risk of bias based on limited reporting of study.
2. Non-significant result.

G.4.1.3 Proactive case finding with palliative care service versus usual care

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Comparator		
Outcome: Length of stay in Hospital (days)									
Campbell (2004)	Cohort study	Very serious[1]	Not serious	N/A	Not serious	26	26	MD -4.70 (-8.87, -0.53)	Low
Outcome Length of stay in ICU days									
Campbell (2004)	Cohort study	Very serious[1]	Not serious	N/A	Not serious	26	26	MD -3.30 (-5.46, -1.14)	Low
Outcome: Reason for discharge (mortality)									
Campbell (2004)	Cohort study	Very serious[1]	Not serious	N/A	Very serious[3]	17/26 (53.8%)	14/26 (65.4%)	RR 0.82 (0.52, 1.29)	Very low
Outcome: Mean length of time (days) from admission until do not resuscitate goals were established									
Campbell (2004)	Cohort study	Very serious[1]	Not serious	N/A	Serious[2]	26	19	MD -1.20 (-3.49, 1.09)	Very low
Outcome: Mean length of stay from establishment of do not resuscitate goals until discharge									
Campbell (2004)	Cohort study	Very serious[1]	Not serious	N/A	Serious[2]	26	19	MD -1.50 (-6.37, 3.37)	Very low
Outcome: Measure of ICU workload (Therapeutic Intervention after DNR-1Scoring System) TISS before DNR-1									
Campbell (2004)	Cohort study	Very serious[1]	Not serious	N/A	Serious[2]	26	19	MD -2.79 (-6.16, 0.58)	Very low
Outcome: Measure of ICU workload TISS after DNR-1									
Campbell (2004)	Cohort study	Very serious[1]	Not serious	N/A	Not serious	26	19	MD -8.24 (-12.84, -3.64)	Low

1. Non-randomised study; high risk of bias based on limited reporting of study.
2. Non-significant result.
3. 95% CI crosses two lines of a defined MID interval

G.4.1.4 Specialist medical and mental health unit versus usual care

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Comparator		
Outcome: Mean difference in MMSE improvement (>2 points) at 90 days									
Goldberg (2013)	RCT	Not serious	Serious[1]	N/A	Very serious[3]	52/163 (32%)	63/167[a] (38%)	RR 0.88 (0.56, 1.37)[b]	Very low
Outcome: Physical disability (Barthel Index) at 90 days									
Goldberg (2013)	RCT	Not serious	Serious[1]	N/A	Serious[2]	187	184	MD -0.1 (-1.1, 0.8)[b]	Low
Outcome: Quality of life (DEMQOL/ 108) at 90 days									
Goldberg (2013)	RCT	Not serious	Serious[1]	N/A	Serious[2]	110	112	MD 0.7 (-2.8, 4.1)[b]	Low
Outcome: Quality of life (DEMQOL proxy/ 124) at 90 days									
Goldberg (2013)	RCT	Not serious	Serious[1]	N/A	Serious[2]	150	138	MD -0.4 (-4.6, 3.8)[b]	Low
Outcome: Quality of life (EQ-5D/1.0 self completed) at 90 days									
Goldberg (2013)	RCT	Not serious	Serious[1]	N/A	Serious[2]	110	112	MD 0.00 (-0.09, 0.09)[b]	Low
Outcome: Quality of life (EQ5D/ 1.0 proxy completed) at 90 days									
Goldberg (2013)	RCT	Not serious	Serious[1]	N/A	Serious[2]	150	138	MD -0.07 (-0.15, 0.00)[b]	Low
Outcome: General health measure (London handicap scale) at 90 days									
Goldberg (2013)	RCT	Not serious	Serious[1]	N/A	Serious[2]	128	123	MD 0.5 (-5.2, 6.2)[b]	Low
Outcome: Number returning home from hospital at 90 days									
Goldberg (2013)	RCT	Not serious	Serious[1]	N/A	Not serious	228/310 (74%)	202/290 (70%)	RR 1.06 (0.95, 1.17)	Moderate
Outcome: Overall mortality at 90 days									

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Comparator		
Goldberg (2013)	RCT	Not serious	Serious[1]	N/A	Serious[4]	68/310 (22%)	71/290 (25%)	RR 0.89 (0.67, 1.19)	Low
Outcome: Readmissions at 90 days									
Goldberg (2013)	RCT	Not serious	Serious[1]	N/A	Serious[4]	99/310 (32%)	101/290 (35%)	RR 0.92 (0.73, 1.15)	Low
Outcome: Carer strain (carer strain Index) at 90 days									
Goldberg (2013)	RCT	Not serious	Serious[1]	N/A	Serious[2]	133	120	MD 0.27 (-0.49, 1.04)[b]	Low

1. Population was mixed delirium/dementia.
2. Non-significant result.
3. 95% CI crosses two lines of a defined MID interval
4. 95% CI crosses one line of a defined MID interval
a. Corrected a numerical typo in published study.
b. Adjusted for age, sex, residence and baseline scores, using multiply imputed data.

G.4.1.5 Follow-up individualised care plan versus usual care

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Comparator		
Outcome: Early ER re-hospitalisation rate (pre- post intervention)									
Villars (2013)	Before/after	Very serious[1]	Not serious	N/A	Very serious[2]	13/168[a] (7.47%)	33/390[a] (8.39%)	RR 0.91 (0.49, 1.69)	Very low
Outcome: Early re- hospitalisation rate in any ward (pre-post intervention)									
Villars (2013)	Before/after	Very serious[1]	Not serious	N/A	Serious[3]	22/168[a] (13.19%)	63/390[a] (16.07%)	RR 0.81 (0.52, 1.23)	Very low
Outcome: ER re-hospital rate at 3 months follow up (pre-post intervention)									

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Comparator		
Villars (2013)	Before/after	Very serious[1]	Not serious	N/A	Serious[3]	39/168[a] (23.58%)	113/390a (28.98%)	RR 0.80 (0.58, 1.09)	Very low
Outcome: Re-hospitalisation in any ward at 3 months follow up (pre-post intervention)									
Villars (2013)	Before/after	Very serious[1]	Not serious	N/A	Serious[3]	21/168[a] (12.70%)	64/390[a] (16.39%)	RR 0.76 (0.48, 1.21)	Very low

1. Selective reporting and limited outcomes (non-randomised study).
2. 95% CI crosses two lines of a defined MID interval
3. 95% CI crosses one line of a defined MID interval

a. Calculations based on percentages reported in published paper.

G.5 Care setting transitions

G.5.1 Managing the transition between different settings for people living with dementia

- What are the most effective ways of managing the transition between different settings (home, care home, hospital, and respite) for people living with dementia?

G.5.1.1 Interventions for people living with dementia

Way-finding interventions

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
Agitation (Pittsburgh Agitation Scale) – lower numbers favour intervention							
1 (McGilton 2003)	Serious[1]	N/A	Not serious	Serious[2]	32	MD 0.28 (-0.44, 1.00)	Low
Spatial orientation (Abilities Assessment Instrument – Spatial Orientation Subscale) – higher numbers favour intervention							
1 (McGilton 2003)	Serious[1]	N/A	Not serious	Serious[2]	32	MD 0.90 (-0.67, 2.47)	Low

1. Lack of blinding (participants and assessors) and allocation concealment
2. Non-significant result

G.5.1.2 Interventions for carers

New York University Caregiver Intervention

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
Carer burden (Zarit Burden Index) – lower numbers favour intervention							
1 (Gaugler 2011)	Serious[1]	N/A	Serious[2]	Serious[3]	406	MD -0.77 (-2.81s, 1.27)	Very low
Carer depression (Geriatric Depression Scale) – lower numbers favour intervention							
1 (Gaugler 2011)	Serious[1]	N/A	Serious[2]	Not serious	406	MD -1.71 (-3.02, -0.40)	Low

1. Lack of blinding (participants)
2. Only outcomes related to carers are reported, not people living with dementia
3. Non-significant result

G.5.1.3 Residential Care Transition Module

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
Carer burden (Zarit Burden Index) – lower numbers favour intervention							
1 (Gaugler 2015)	Serious[1]	N/A	Serious[2]	Serious[3]	36	MD -2.86 (-6.71, 0.99)	Very low
Carer stress (Perceived Stress Scale) – lower numbers favour intervention							
1 (Gaugler 2015)	Serious[1]	N/A	Serious[2]	Serious[3]	36	MD -5.08 (-10.32, 0.16)	Very low
Carer depression (Center for Epidemiologic Studies-Depression Scale) – lower numbers favour intervention							
1 (Gaugler 2015)	Serious[1]	N/A	Serious[2]	Serious[3]	36	MD -5.00 (-12.01, 2.01)	Very low
Carer satisfaction with facility (Likert scale) – higher numbers favour intervention							
1 (Gaugler 2015)	Serious[1]	N/A	Serious[2]	Serious[3]	36	MD 0.24 (-0.06, 0.54)	Very low
Carer satisfaction with role (Family Caregiver Perception Role Scale) – higher numbers favour intervention							
1 (Gaugler 2015)	Serious[1]	N/A	Serious[2]	Serious[3]	36	MD -0.09 (-0.80, 0.62)	Very low

1. Lack of blinding (participants and assessors)
2. Only outcomes related to carers are reported, not people living with dementia
3. Non-significant result

G.5.1.4 FITT-NH (Family Intervention: Telephone Tracking-Nursing Home)

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
Carer burden (Zarit Burden Index) – lower numbers favour intervention							
1 (Davies 2011)	Serious[1]	N/A	Serious[2]	Serious[3]	46	MD -5.07 (-12.13, 1.99)	Very low
Carer depression (Center for Epidemiology Studies Depression Scale) – lower numbers favour intervention							
1 (Davies 2011)	Serious[1]	N/A	Serious[2]	Serious[3]	46	MD 0.29 (-5.62, 6.20)	Very low
Carer satisfaction with facility (Likert scale) – higher numbers favour intervention							
1 (Davies 2011)	Serious[1]	N/A	Serious[2]	Serious[3]	46	MD 0.31 (-0.05, 0.67)	Very low

1. Lack of blinding (participants and assessors) and allocation concealment
2. Only outcomes related to carers are reported, not people living with dementia
3. Non-significant result

G.6 Modifying risk factors for dementia progression

G.6.1 Risk factors for dementia progression

- What effect does modifying risk factors have on slowing the progression of dementia?

G.6.1.1 Antidiabetic drugs versus placebo

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Control	Summary of results	
Cognition – ADAS-cog (6 months) - lower numbers favour antidiabetic drugs									
2 (Gold 2010, Risner 2006)	RCT	Serious[1]	Not serious	Not serious	Serious[2]	512	252	MD -0.42 (-1.35, 0.51)	Low
Cognition – MMSE (6 months) - higher numbers favour antidiabetic drugs									
1 (Gold 2010)	RCT	Serious[1]	Not serious	N/A	Very serious[2,3]	260	131	Non-significant (MD not reported)	Very low
Clinical Global Assessment – CIBIC+ (6 months) - lower numbers favour antidiabetic drugs									
1 (Gold 2010)	RCT	Serious[1]	Not serious	N/A	Serious[2]	260	131	MD -0.05 (-0.27, 0.17)	Low
Behavioural symptoms – NPI (6 months) - lower numbers favour antidiabetic drugs									
1 (Gold 2010)	RCT	Serious[1]	Not serious	N/A	Very serious[2,3]	260	131	Non-significant (MD not reported)	Very low
Any adverse event (6 months)									
2 (Gold 2010, Risner 2006)	RCT	Serious[1]	Not serious	Not serious	Serious[4]	594	288	RR 0.97 (0.80,1.16)	Low
Serious adverse events (6 months)									
2 (Gold 2010, Risner 2006)	RCT	Serious[1]	Not serious	Not serious	Very serious[5]	594	288	RR 0.91 (0.50, 1.64)	Very low
Adverse events leading to discontinuation (6 months)									
1 (Gold 2010)	RCT	Serious[1]	Not serious	Not serious	Very serious[5]	331	164	RR 0.99 (0.43, 2.27)	Very low

1. Participants were allowed to take other medications (such as antipsychotics, antidepressants and vitamin E supplements) which may have had an impact the outcome measure of interest; however, it was not reported what proportions of participants in each group took these medications.
2. Non-significant result.

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Control	Summary of results	

3. Mean difference and measures of dispersion not reported.
4. 95% CI crosses two lines of a defined MID interval.
5. 95% CI crosses one line of a defined MID interval.

G.6.1.2 NSAIDs versus placebo

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Control	Summary of results	
Cognition – ADAS-cog (6 months) – lower numbers favour NSAIDs									
4	RCT	Serious[1]	Not serious	Not serious	Serious[2]	1,097	918	MD -0.00 (-0.53, 0.53)	Low
Cognition – ADAS-cog (12 months) – lower numbers favour NSAIDs									
7	RCT	Serious[1]	Not serious	Serious[3]	Serious[2]	1,743	1,541	MD -0.25 (-1.89, 1.40)	Low
Cognition – MMSE (6 months) – higher numbers favour NSAIDs									
6	RCT	Serious[1]	Not serious	Not serious	Serious[2]	292	165	MD -0.33 (-0.81, 0.15)	Low
Cognition – MMSE (12 months) – higher numbers favour NSAIDs									
6	RCT	Very serious[1,4]	Not serious	Not serious	Serious[2]	1,375	1,231	MD -0.22 (-0.47, 0.03)	Very low
Functional ability – ADCS-ADL (6 months) – higher numbers favour NSAIDs									
1 (Green 2009)	RCT	Serious[1]	Not serious	N/A	Serious[2]	751	725	MD -0.41 (-1.20, 0.38)	Low
Functional ability – ADCS-ADL (12 months) – higher numbers favour NSAIDs									
4	RCT	Serious[1]	Not serious	Serious[3]	Not serious	1,350	1,321	MD 1.60 (0.31, 2.90)	Low
Functional ability – ADCS-ADL, IDDD & BADLS (12 months: SMD) – higher numbers favour NSAIDs									
7	RCT	Very serious[1,4]	Not serious	Not serious	Not serious	1,512	1,477	SMD 0.10 (0.02, 0.17)	Moderate
Global assessment – CIBIC+ (6 months) – lower numbers favour NSAIDs									
2	RCT	Serious[1]	Not serious	Not serious	Serious[2]	296	158	MD 0.06 (-0.12, 0.24)	Low
Global assessment – CIBIC+ & CGIC (6 months: SMD) – lower numbers favour NSAIDs									
3	RCT	Serious[1]	Not serious	Not serious	Serious[5]	313	172	SMD 0.04 (-0.15, 0.23)	Low
Global assessment – CIBIC+ (12 months) – lower numbers favour NSAIDs									
4	RCT	Serious[1]	Not serious	Not serious	Serious[2]	668	528	MD 0.04 (-0.08, 0.16)	Low

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Control	Summary of results	
Behavioural symptoms: NPI (6 months) – lower numbers favour NSAIDs									
2	RCT	Serious[1]	Not serious	Not serious	Serious[2]	787	750	MD -0.01 (-0.91, 0.89)	Low
Behavioural symptoms: NPI & Behave-AD (6 months: SMD) – lower numbers favour NSAIDs									
3	RCT	Serious[1]	Not serious	Not serious	Not serious	1,062	885	SMD 0.03 (-0.06, 0.12)	Moderate
Behavioural symptoms: NPI (12 months) – lower numbers favour NSAIDs									
4	RCT	Serious[1]	Not serious	Not serious	Serious[2]	1,061	1,012	MD -0.32 -0.95, 0.31)	Low
Behavioural symptoms: NPI & Behave-AD (12 months: SMD) – lower numbers favour NSAIDs									
5	RCT	Serious[1]	Not serious	Serious[3]	Not serious	1,337	1,147	SMD 0.02 (-0.06, 0.10)	Low
Dementia severity: CDR-SB (12 months) – lower numbers favour NSAIDs									
5	RCT	Serious[1]	Not serious	Serious[3]	Serious[2]	1,424	1,379	MD 0.03 (-0.15, 0.21)	Very low
Quality of life: QoL-AD (12 months)									
2	RCT	Serious[1]	Not serious	Not serious	Serious[2]	810	775	MD 0.31 (-0.26, 0.88)	Low
Any adverse events (12 months)									
4	RCT	Serious[1]	Not serious	Not serious	Not serious	1,561	1,373	RR 1.03 (1.00, 1.07)	Moderate
Serious adverse events (12 months)									
6	RCT	Very serious[1,4]	Not serious	Not serious	Serious[6]	1,913	1,673	RR 1.16 (1.02, 1.31)	Very low
Adverse events leading to discontinuation (12 months)									
6	RCT	Serious[1]	Not serious	Not serious	Serious[6]	1,867	1,666	RR 1.44 (1.20, 1.73)	Low
Mortality (12 months)									
4	RCT	Serious[1]	Not serious	Not serious	Very serious[7]	690	458	RR 1.63 (0.71, 3.71)	Very low

1. Participants were allowed to take other medications (such as antipsychotics, antidepressants and vitamin E supplements) which may have had an impact the outcome measure of interest; however, it was not reported what proportions of participants in each group took these medications.
2. Non-significant result.
3. I^2>40%
4. Assessors not blinded to group allocation
5. Confidence interval crosses one line of a defined minimum clinically important difference (SMDs of -0.2 and 0.2)
6. 95% CI crosses one line of a defined MID interval.
7. 95% CI crosses two lines of a defined MID interval.

G.6.1.3 Statins versus placebo

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Control	Summary of results	
Cognition – ADAS-cog (6 months) – lower numbers favour NSAIDs									
4	RCT	Serious[1]	Not serious	Not serious	Serious[2]	551	516	MD -0.08 (-0.85, 0.70)	Low
Cognition – ADAS-cog (12 months) – lower numbers favour NSAIDs									
2	RCT	Serious[1]	Not serious	Not serious	Serious[2]	440	480	MD -0.12 (-1.04, 0.80)	Low
Cognition – MMSE (6 months) – higher numbers favour NSAIDs									
4	RCT	Serious[1]	Not serious	Serious[3]	Serious[2]	523	561	MD 0.48 (-0.12, 1.08)	Very low
Cognition – MMSE (12 months) – higher numbers favour NSAIDs									
3	RCT	Serious[1]	Not serious	Serious[3]	Serious[2]	472	511	MD 0.42 (-0.37, 1.20)	Very low
Behavioural symptoms – NPI (6 months) – lower numbers favour NSAIDs									
3	RCT	Serious[1]	Not serious	Serious[3]	Serious[2]	498	541	MD -1.59 (-3.47, 0.29)	Very low
Behavioural symptoms – NPI (12 months) – lower numbers favour NSAIDs									
3	RCT	Serious[1]	Not serious	Serious[3]	Serious[2]	472	511	MD -1.64 (-3.45, 0.18)	Very low
Any adverse events (12 months)									
2	RCT	Serious[1]	Not serious	Serious[3]	Very serious[4]	396	527	RR 1.71 (0.39, 7.60)	Very low
Serious adverse events (12 months)									
3	RCT	Serious[1]	Not serious	Not serious	Serious[5]	518	527	RR 0.96 (0.77, 1.19)	Low
Adverse events leading to discontinuation (12 months)									
1 (Feldman 2010)	RCT	Serious[1]	Not serious	N/A	Not serious	314	325	RR 7.45 (2.96, 18.75)	Moderate
Mortality (12 months)									
2	RCT	Serious[1]	Not serious	Serious[1]	Very serious[3]	518	527	RR 0.94 (0.34, 2.59)	Very low

1. Participants were allowed to take other medications (such as antipsychotics, antidepressants and vitamin E supplements) which may have had an impact the outcome measure of interest; however, it was not reported what proportions of participants in each group took these medications.
2. Non-significant result
3. I^2>40%
4. 95% CI crosses two lines of a defined MID interval.
5. 95% CI crosses one line of a defined MID interval.

G.6.1.4 Antihypertensive drugs

Calcium-channel blocker versus placebo

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Control	Summary of results	
Cognition – ADAS-cog (6 months) – lower numbers favour calcium-channel blocker									
1 (Morich 2012)	RCT	Serious[1]	Not serious	N/A	Serious[2]	958	484	MD -0.45 (-1.09, 0.20)	Low
Cognition – MMSE (6 months) – higher numbers favour calcium-channel blocker									
1 (Morich 2012)	RCT	Serious[1]	Not serious	N/A	Not serious	958	484	MD 0.35 (0.13, 0.56)	Moderate
Cognition – MMSE (12 months) – higher numbers favour calcium-channel blocker									
1 (Pantoni 2005)	RCT	Serious[1]	Not serious	N/A	Serious[2]	94	55	MD 0.60 (-1.64, 2.84)	Low
Global assessment – CGI, global improvement (6 months) – lower numbers favour calcium-channel blocker									
1 (Morich 2012)	RCT	Serious[1]	Not serious	N/A	Serious[2]	958	484	RR 0.04 (-0.07, 0.14)	Low
Any adverse events (6 months)									
1 (Morich 2012)	RCT	Serious[1]	Not serious	N/A	Not serious	1,086	550	RR 1.01 (0.95, 1.08)	Moderate
Serious adverse events (6 months)									
1 (Morich 2012)	RCT	Serious[1]	Not serious	N/A	Not serious	1,086	550	RR 2.25 (1.32, 3.83)	Moderate
Adverse events leading to discontinuation (6 months)									
1 (Morich 2012)	RCT	Serious[1]	Not serious	N/A	Very serious[3]	1,086	550	RR 1.17 (0.77, 1.77)	Very low

1. Participants were allowed to take other medications (such as antipsychotics, antidepressants and vitamin E supplements) which may have had an impact the outcome measure of interest; however, it was not reported what proportions of participants in each group took these medications.
2. Non-significant result
3. 95% CI crosses two lines of a defined MID interval.

G.6.1.5 Angiotensin II receptor antagonist versus calcium-channel blocker

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Angiotensin II receptor antagonist	Calcium channel blocker	Summary of results	
Cognition – MMSE (6 months) – higher numbers favour angiotensin II receptor antagonist									

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Angiotensin II receptor antagonist	Calcium channel blocker	Summary of results	
1 (Kume 2012)	RCT	Not serious	Not serious	N/A	Serious[1]	10	10	MD 1.3 (-1.80, 4.40)	Moderate
Cognition – ADAS-cog (6 months) – lower numbers favour angiotensin II receptor antagonist									
1 (Kume 2012)	RCT	Not serious	Not serious	N/A	Serious[1]	10	10	MD -4.2 (-9.42, 1.02)	Moderate
1. Non-significant result									

G.6.1.6 Brain-penetrating angiotensin converting enzyme (ACE) inhibitor versus calcium-channel blocker

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	ACE inhibitor	Calcium channel blocker	Summary of results	
Cognition – MMSE (12 months) – higher numbers favour ACE inhibitor									
1 (Ohrui 2004)	RCT	Serious[1]	Not serious	N/A	Not serious	51	57	MD 4.3 (4.22, 4.38)	Moderate
1. Authors do not report whether patients or assessors were blinded to group allocations									

G.6.1.7 Non-brain-penetrating ACE inhibitor versus calcium-channel blocker

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	ACE inhibitor	Calcium channel blocker	Summary of results	
Cognition – MMSE (12 months) – higher numbers favour ACE inhibitor									
1 (Ohrui 2004)	RCT	Serious[1]	Not serious	N/A	Not serious	51	57	MD 0.3 (0.19, 0.38)	Moderate
1. Authors do not report whether patients or assessors were blinded to group allocations									

G.7 Cholinesterase inhibitors and memantine for dementia

G.7.1 Cholinesterase inhibitors and memantine for people living with Alzheimer's disease

- Who should start and review the following pharmacological interventions: (donepezil, galantamine, rivastigmine, memantine) for people with Alzheimer's disease and how should a review be carried out?

Prescribing donepezil

Quality assessment						No of patients		Effect size (95% CI)	Quality
No of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Geriatric Psychiatrist (GERO)**	**Primary care physician (MED)**		
Clinical outcome (including cognitive, functional & behavioural ability)									
Outcome 1: Mean Clinical Dementia Rating (CDR) scores at 1 year follow up									
Aupperle (2000)	Retrospective cohort study	Very serious[1]	N/A	Not serious	Not serious	26	31	MD 0.70 (0.36, 1.04)	Low
Concordance & compliance									
Outcome 1: Provider practices- prescription of donepezil at 1 year follow up									
Aupperle (2000)	Retrospective cohort	Very serious[1]	N/A	Not serious	Not serious	20/26	11/31	RR 0.46 (0.27, 0.78)	Low
Access to health and social care support									
Outcome 1: Service usage (past 6 months): Number of people receiving hospitalisation									
Aupperle (2000)	Retrospective cohort study	Very serious[1]	N/A	Not serious	Serious[2]	4/26	12/31	RR 2.52 (0.92, 6.87)	Very low
Outcome 2: Service usage (past 6 months): Number of people receiving home health aide									
Aupperle (2000)	Retrospective cohort study	Very serious[1]	N/A	Not serious	Serious[2]	5/26	14/31	RR 2.35 (0.98, 5.65)	Very low
Outcome 3: Service usage (past 6 months): Number of people attending dementia day program									

Quality assessment						No of patients		Effect size (95% CI)	Quality
No of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Geriatric Psychiatrist (GERO)**	**Primary care physician (MED)**		
Aupperle (2000)	Retrospective cohort study	Very serious[1]	N/A	Not serious	Very serious[3]	7/26	5/31	RR 0.60 (0.22, 1.67)	Very low
Patient and carer experience and satisfaction									
Outcome 1: Carer distress rating (Zarit Burden Interview) at 1 year follow up									
Aupperle (2000)	Retrospective cohort study	Very serious[1]	N/A	Not serious	Serious[4]	26	31	MD 2.40 (-4.16, 8.96)	Very low

1. Included study at high risk of bias
2. 95% CI crosses one line of a defined MID interval
3. 95% CI crosses two lines of a defined MID interval
4. Non-significant result

Reviewing donepezil

Quality assessment						No of patients		Effect size (95% CI)	Quality
No of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Not receiving advisory service (Non DOCS)**	**Receiving advisory service (DOCS)**		
Concordance & compliance									
Outcome 1: Medication persistence rate: Mean duration of donepezil treatment									
Watanabe (2012)	Before and after study	Very serious[1]	N/A	Very serious[2]	Not serious	59	52	MD 130.4 (58.02, 202.8)	Very low

Quality assessment						No of patients		Effect size (95% CI)	
No of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Not receiving advisory service (Non DOCS)**	**Receiving advisory service (DOCS)**		**Quality**
Outcome 2: Medication persistence rate: Use of donepezil at 1 year follow up									
Watanabe (2012)	Before and after study	Very serious[1]	N/A	Very serious[2]	Serious[3]	29/59	38/52	RR 1.49 (1.09, 2.02)	Very low
Patient and carer experience and satisfaction									
Outcome 1: Average level of carer understanding at 4 week follow up									
Watanabe (2012)	Before and after study	Very serious[1]	N/A	Very serious[2]	Not serious	26	31	MD 3.20 (2.70, 3.70)	Very low

1. Downgraded due to observational study. Short follow up period (4 weeks) for outcomes, validation of scale used for survey of understanding not clearly reported
2. Non UK setting and indirect setting for advisory consultation service
3. 95% CI crosses one line of a defined MID interval

G.7.2 Cholinesterase inhibitors and memantine in Alzheimer's disease

- How effective is the co-prescription of cholinesterase inhibitors and memantine for the treatment of Alzheimer's disease?
- When should treatment with donepezil, galantamine, rivastigmine, memantine be withdrawn for people with Alzheimer's disease?

G.7.2.1 Any cholinesterase inhibitor plus memantine versus any cholinesterase inhibitor plus placebo

Full population

Quality assessment						No of patients		Effect estimate	
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Combination therapy	AChEI monotherapy	Effect size (95% CI)	Quality
Cognition: (ADAS-cog) lower values favour intervention									
Dysken 2014; Porsteinsson 2008	RCT	Not serious	Not serious	Not serious	Serious[2]	356	353	MD -0.63 (-2.13, 0.87)	Moderate
Cognition: (MMSE) higher values favour intervention									
Dysken 2014; Howard 2012[a] Porsteinsson 2008	RCT	Not serious	Not serious	Not serious	Serious[2]	410	392	MD 0.14 (-0.47, 0.75)	Moderate
Activities of daily living (ADCS-ADL/BADLS) higher values favour intervention									
Grossberg 2013; Howard 2012[a]; Tariot 2004; Dysken 2014; Porsteinsson 2008	RCT	Not serious	Not serious	Not serious	Not serious	943	932	SMD 0.10 (0.01, 0.19)	High
Global functioning (CIBIC plus) lower values favour intervention									
Grossberg 2013; Tariot 2004; Porsteinsson 2008	RCT	Not serious	Not serious	Serious[1]	Not serious	745	738	MD -0.20 (-0.36, -0.04)	Moderate
Behavioural and psychological symptoms (NPI) lower values favour intervention									
Grossberg 2013; Howard 2012[a]; Tariot	RCT	Not serious	Not serious	Not serious	Not serious	923	913	MD -1.91 (-3.16, -0.65)	High

Quality assessment						No of patients		Effect estimate	
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Combination therapy	AChEI monotherapy	Effect size (95% CI)	Quality
2004; Dysken 2014; Porsteinsson 2008									
Care dependency (Behaviour rating scale for geriatric patients- care dependency subscale) lower values favour intervention									
Tariot 2004	RCT	Not serious	Not serious	N/A	Not serious	185	179	MD -1.50 (-2.54, -0.46)	High
Severe impairment battery (SIB)									
Grossberg 2013; Tariot 2004	RCT	Not serious	Not serious	Serious[1]	Serious[2]	530	523	MD 1.22 (-1.15, 3.59)	Low
Verbal fluency test (VFT) higher values favour intervention									
Grossberg 2013		Not serious	Not serious	N/A	Not serious	330	326	MD 0.60 (0.19, 1.01)	High
Health related quality of life (DEMQOL) higher values favour intervention									
Howard 2012[a]	RCT	Not serious	Not serious	N/A	Serious[2]	58	55	MD -2.00 (-6.44, 2.44)	Moderate
Global health questionnaire (GHQ) higher values favour intervention									
Howard 2012[a]	RCT	Not serious	Not serious	N/A	Serious[2]	54	45	MD 0.13 (-0.87, 1.13)	Moderate
Total number of adverse events: lower values favour intervention									
Grossberg 2013; Tariot 2004 Dysken 2014[b]	RCT	Not serious	Not serious	Not serious	Not serious	698	688	RR 1.00 (0.93, 1.09)	High
Number of serious adverse events: lower values favour intervention									
Grossberg 2013; Howard 2012; Dysken 2014[b] Porsteinsson 2008	RCT	Not serious	Not serious	Not serious	Serious[3]	789	766	RR 0.95 (0.76, 1.19)	Moderate
Number of discontinuations to adverse events: lower values favour intervention									
Grossberg 2013;	RCT	Not serious	Not serious	Serious[1]	Very serious[4]	760	752	RR 0.92 (0.49, 1.71)	Low

Quality assessment						No of patients		Effect estimate	
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Combination therapy	AChEI monotherapy	Effect size (95% CI)	Quality
Tariot 2004; Porsteinsson 2008									
Mortality: lower values favour intervention									
Grossberg 2013; Howard 2012; Dysken 2014; Porsteinsson 2008	RCT	Not serious	Not serious	Not serious	Serious[3]	789	776	RR 1.14 (0.80, 1.62)	Moderate
Caregiver activity survey (CAS): higher values favour intervention									
Dysken 2014	RCT	Not serious	Not serious	N/A	Serious[2]	142	140	MD 0.38 (-1.80, 2.56)	Moderate
Entry to care home: lower numbers favour intervention									
Howard 2012	RCT	Not serious	Not serious	N/A	Serious[2]	73	73	HR 1.22 (0.78, 1.90)	Moderate

1. I^2 >40%
2. Non-significant result
3. 95% CI crosses one line of a defined MID interval
4. 95% CI crosses two lines of a defined MID interval

a: extracted from additional data (see appendix E)
b: Number of adverse events authors attributed to study medication

Mild to moderate

Quality assessment						No of patients		Effect estimate	
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Combination therapy	AChEI monotherapy	Effect size (95% CI)	Quality
Cognition: (ADAS-cog) lower values favour intervention									
Dysken 2014; Porsteinsson 2008	RCT	Not serious	Not serious	Not serious	Serious[1]	356	353	MD -0.63 (-2.13, 0.87)	Moderate

Quality assessment						No of patients		Effect estimate	
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Combination therapy	AChEI monotherapy	Effect size (95% CI)	Quality
Cognition: (MMSE) higher values favour intervention									
Dysken 2014; Howard 2012[a]	RCT	Not serious	Not serious	Not serious	Serious[1]	352	338	MD 0.11 (-0.57, 0.78)	Moderate
Activities of daily living (ADCS-ADL/BADLS) higher values favour intervention									
Dysken 2014; Porsteinsson 2008	RCT	Not serious	Not serious	Not serious	Serious[2]	356	353	SMD 0.05 (-0.10, 0.20)	Moderate
Global functioning (CIBIC plus) lower values favour intervention									
Porsteinsson 2008	RCT	Not serious	Not serious	N/A	Serious[1]	214	213	MD -0.04 (-0.23, 0.15)	Moderate
Behavioural and psychological symptoms (NPI) lower values favour intervention									
Dysken 2014; Porsteinsson 2008	RCT	Not serious	Not serious	Not serious	Serious[1]	354	349	MD -0.04 (-2.01, 1.92)	Moderate
Health related quality of life (DEMQOL) higher values favour intervention									
Howard 2012[a]	RCT	Not serious	Not serious	N/A	Serious[1]	58	55	MD -2.00 (-6.44, 2.44)	Moderate
Total number of adverse events: lower values favour intervention									
Dysken 2014[b]	RCT	Not serious	Not serious	N/A	Very serious[3]	155	152	RR 1.18 (0.72, 1.94)	Low
Number of serious adverse events: lower values favour intervention									
Dysken 2014[b] Porsteinsson 2008	RCT	Not serious	Not serious	Not serious	Very serious[3]	372	368	RR 0.91 (0.62, 1.33)	Low
Number of discontinuations to adverse events: lower values favour intervention									
Porsteinsson 2008	RCT	Not serious	Not serious	N/A	Very serious[3]	217	216	RR 0.76 (0.38, 1.53)	Low
Mortality: lower values favour intervention									
Dysken 2014;	RCT	Not serious	Not serious	Not serious	Serious[2]	372	368	RR 1.25 (0.83, 1.87)	Moderate

Quality assessment						No of patients		Effect estimate	
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Combination therapy	AChEI monotherapy	Effect size (95% CI)	Quality
Porsteinsson 2008									
Caregiver activity survey (CAS) higher values favour intervention									
Dysken 2014	RCT	Not serious	Not serious	N/A	Serious[1]	142	140	MD 0.38 (-1.80, 2.56)	Moderate

1. Non-significant result
2. 95% CI crosses one line of a defined MID interval
3. 95% CI crosses two lines of a defined MID interval

a: extracted from additional data (see appendix E)
b: Number of adverse events authors attributed to study medication

Moderate to severe

Quality assessment						No of patients		Effect estimate	
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Combination therapy	AChEI monotherapy	Effect size (95% CI)	Quality
Cognition: (MMSE) higher values favour intervention									
Howard 2012[a]	RCT	Not serious	Not serious	N/A	Serious[2]	58	54	MD 0.27 (-1.13, 1.67)	Moderate
Activities of daily living (ADCS-ADL/BADLS) higher values favour intervention									
Grossberg 2013; Howard 2012[a]; Tariot 2004	RCT	Not serious	Not serious	Not serious	Serious[3]	587	579	SMD 0.13 (0.01, 0.24)	Moderate
Global functioning (CIBIC plus) lower values favour intervention									
Grossberg 2013; Tariot 2004	RCT	Not serious	Not serious	Serious[1]	Not serious	531	525	MD -0.28 (-0.41, -0.14)	Moderate
Behavioural and psychological symptoms (NPI) lower values favour intervention									

Quality assessment						No of patients		Effect estimate	
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Combination therapy	AChEI monotherapy	Effect size (95% CI)	Quality
Grossberg 2013; Howard 2012[a]; Tariot 2004	RCT	Not serious	Not serious	Not serious	Not serious	569	564	MD -3.19 (-4.83, -1.56)	High
Care dependency (Behaviour rating scale for geriatric patients- care dependency subscale) lower values favour intervention									
Tariot 2004	RCT	Not serious	Not serious	N/A	Not serious	185	179	MD -1.50 (-2.54, -0.46)	High
Severe impairment battery (SIB): higher values favour intervention									
Grossberg 2013; Tariot 2004	RCT	Not serious	Not serious	Serious[1]	Serious[2]	530	523	MD 1.22 (-1.15, 3.59)	Low
Verbal fluency test (VFT) higher values favour intervention									
Grossberg 2013		Not serious	Not serious	N/A	Not serious	330	326	MD 0.60 (0.19, 1.01)	High
Health related quality of life (DEMQOL) higher values favour intervention									
Howard 2012[a]	RCT	Not serious	Not serious	N/A	Serious[2]	58	55	MD -2.00 (-6.44, 2.44)	Moderate
Global health questionnaire (GHQ) higher values favour intervention									
Howard 2012[a]	RCT	Not serious	Not serious	N/A	Serious[2]	54	45	MD 0.13 (-0.87, 1.13)	Moderate
Total number of adverse events: lower values favour intervention									
Grossberg 2013; Tariot 2004	RCT	Not serious	Not serious	Not serious	Not serious	372	370	RR 0.99 (0.92, 1.08)	High
Number of serious adverse events: lower values favour intervention									
Grossberg 2013; Howard 2012;	RCT	Not serious	Not serious	Serious[1]	Very serious[4]	417	408	RR 0.98 (0.76, 1.28)	Very low
Number of discontinuations to adverse events: lower values favour intervention									
Grossberg 2013; Tariot 2004; Porsteinsson 2008	RCT	Not serious	Not serious	Serious[1]	Very serious[4]	543	536	RR 0.99 (0.38, 2.58)	Very low

Quality assessment						No of patients		Effect estimate	
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Combination therapy	AChEI monotherapy	Effect size (95% CI)	Quality
Mortality: lower values favour intervention									
Grossberg 2013; Howard 2012;	RCT	Not serious	Not serious	Not serious	Very serious[4]	417	408	RR 0.90 (0.45, 1.80)	Low

1. I^2 >40%
2. Non-significant result
3. 95% CI crosses one line of a defined MID interval
4. 95% CI crosses two lines of a defined MID interval

a: extracted from additional data (see appendix E)

Mild only

Quality assessment						No of patients		Effect estimate	
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Combination therapy	AChEI monotherapy	Effect size (95% CI)	Quality
Clinical Global: post-hoc within-trial subgroup analyses (lower values favour intervention)									
Porsteinsson 2008	RCT	Not serious	Not serious	N/A	Very serious[2]	57	64	SMD -0.09 (-0.45, 0.26)	Low
Cognitive Function: post-hoc within-trial subgroup analyses (lower values favour intervention)									
Dysken 2014; Porsteinsson 2008	RCT	Not serious	Not serious	Not serious	Serious[1]	162	153	SMD -0.05 (-0.27, 0.17)	Moderate
Decline in Activities of Daily Living: post-hoc within-trial subgroup analyses (lower values favour intervention)									
Dysken 2014; Porsteinsson 2008	RCT	Not serious	Not serious	Not serious	Serious[1]	162	153	SMD -0.04 (-0.26, 0.19)	Moderate

1. 95% CI crosses one line of a defined MID interval
2. 95% CI crosses two lines of a defined MID interval

a: extracted from additional data (see appendix E)

Moderate only

Quality assessment						No of patients		Effect estimate	
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Combination therapy	AChEI monotherapy	Effect size (95% CI)	Quality
Clinical Global: post-hoc within-trial subgroup analyses (lower values favour intervention)									
Porsteinsson 2008; Tariot 2004	RCT	Not serious	Not serious	Serious[1]	Serious[2]	294	312	SMD -0.17 (-0.35, 0.00)	Low
Cognitive Function: post-hoc within-trial subgroup analyses (lower values favour intervention)									
Dysken 2014; Howard 2012; Porsteinsson 2008 Tariot 2004	RCT	Not serious	Not serious	Not serious	Serious[2]	319	338	SMD -0.23 (-0.39, -0.08)	Moderate
Decline in Activities of Daily Living: post-hoc within-trial subgroup analyses (lower values favour intervention)									
Dysken 2014; Howard 2012; Porsteinsson 2008 Tariot 2004	RCT	Not serious	Not serious	Not serious	Serious[2]	322	341	SMD -0.04 (-0.26, 0.19)	Moderate
NPI (lower values favour intervention)									
Howard 2012	RCT	Not serious	Not serious	N/A	Serious[3]	27	28	MD 0.47 (-10.43, 11.37)	Moderate
DEMQOL (higher values favour intervention)									
Howard 2012	RCT	Not serious	Not serious	N/A	Serious[3]	27	28	MD -4.45 (-11.34, 2.44)	Moderate
GHQ-12 (higher values favour intervention)									
Howard 2012	RCT	Not serious	Not serious	N/A	Serious[3]	24	28	MD 0.31 (-1.32, 1.94)	Moderate

1. I^2 >40%
2. 95% CI crosses one line of a defined MID interval
3. Non-significant result

a: extracted from additional data (see appendix E)

Severe only

Quality assessment						No of patients		Effect estimate	
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Combination therapy	AChEI monotherapy	Effect size (95% CI)	Quality
Clinical Global: post-hoc within-trial subgroup analyses (lower values favour intervention)									
Tariot 2004	RCT	Not serious	Not serious	N/A	Serious[2]	89	72	SMD -0.22 (-0.53, 0.09)	Moderate
Cognitive Function: post-hoc within-trial subgroup analyses (lower values favour intervention)									
Dysken 2014; Howard 2012; Tariot 2004	RCT	Not serious	Not serious	Not serious	Not serious	120	98	SMD -0.57 (-0.84, -0.30)	High
Decline in Activities of Daily Living: post-hoc within-trial subgroup analyses (lower values favour intervention)									
Howard 2012; Tariot 2004	RCT	Not serious	Not serious	Not serious	Serious[2]	120	98	SMD -0.33 (-0.60, -0.06)	Moderate
NPI (lower values favour intervention)									
Howard 2012	RCT	Not serious	Not serious	N/A	Not serious	31	26	MD -10.24 (-20.30, -0.18)	High
DEMQOL (higher values favour intervention)									
Howard 2012	RCT	Not serious	Not serious	N/A	Serious[1]	31	26	MD 0.49 (-6.02, 7.00)	Moderate
GHQ-12 (higher values favour intervention)									
Howard 2012	RCT	Not serious	Not serious	N/A	Serious[1]	30	23	MD -0.10 (-1.32, 1.12)	Moderate

1. Non-significant result
2. 95% CI crosses one line of a defined MID interval

a: extracted from additional data (see appendix E)

G.7.2.2 Any cholinesterase inhibitor plus memantine versus cholinesterase inhibitor monotherapy

Quality assessment						No of patients		Effect estimate	
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Combination therapy	AChEI monotherapy	Effect size (95% CI)	Quality
Cognition: MMSE higher values favour intervention									
Araki 2014; Choi 2011	RCT	Serious[1]	Not serious	Serious[2]	Serious[3]	96	87	MD 0.88 (-1.98, 3.75)	Very low
Cognition: ADAS-cog lower values favour intervention									
Choi 2011	RCT	Serious[1]	Not serious	N/A	Serious[3]	84	74	MD -0.66 (-2.81, 1.49)	Low
Global (Clinical Global Impression- Improvement) lower values favour intervention									
Araki 2014	RCT	Serious[1]	Not serious	N/A	Not serious	12	13	MD -2.60 (-3.44, -1.76)	Moderate
Clock Drawing Test (CDT) higher values favour intervention									
Araki 2014	RCT	Serious[1]	Not serious	N/A	Not serious	12	13	MD 3.59 (1.39, 5.79)	Moderate
Neuropsychiatric (NPI) lower values favour intervention									
Araki 2014	RCT	Serious[1]	Not serious	N/A	Not serious	12	13	MD -23.71 (-32.51, -14.91)	Moderate
Neuropsychiatric (NPI) caregiver administered lower values favour intervention									
Choi 2011	RCT	Serious[1]	Not serious	N/A	Serious[3]	84	74	MD 0.20 (-35.87, 36.27)	Low
Frontal Assessment Battery (FAB) lower values favour intervention									
Choi 2011	RCT	Serious[1]	Not serious	N/A	Serious[3]	84	74	MD -0.20 (-0.93, 0.53)	Low
Clinical Dementia rating (sum of boxes) higher values favour intervention									
Choi 2011	RCT	Serious[1]	Not serious	N/A	Serious[3]	84	74	MD 0.11 (-0.40, 0.62)	Low
Cohen Mansfield Agitation Inventory (CMAI) lower values favour intervention									
Choi 2011	RCT	Serious[1]	Not serious	N/A	Serious[3]	84	74	MD 1.00 (-1.57, 3.57)	Low
Japanese Zarit Burden Interview (J-ZBI) lower values favour intervention									

Choi 2011	RCT	Serious[1]	Not serious	N/A	Not serious	84	74	MD -18.56 (-26.06, -11.06)	Moderate
Any adverse event: lower values favour intervention									
Choi 2011	RCT	Serious[1]	Not serious	N/A	Serious[3]	84	74	MD 1.06 (0.79, 1.41)	Low
Any serious adverse event: lower values favour intervention									
Choi 2011	RCT	Serious[1]	Not serious	N/A	Serious[3]	84	74	MD 1.89 (0.35, 10.03)	Low

1. Not placebo controlled
2. I^2 >40%
3. Non-significant result

G.7.2.3 Any cholinesterase inhibitor plus memantine versus memantine plus placebo

Quality assessment						No of patients		Effect estimate	
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Combination therapy	AChEI monotherapy	Effect size (95% CI)	Quality
Cognition: MMSE higher values favour intervention									
Shao 2015	RCT	Serious[1]	Not serious	Not serious[4]	Serious[3]	66	22	MD 0.54 (-0.30, 1.38)	Low
Activities of Daily living (ADCS-ADL) higher values favour intervention									
Shao 2015	RCT	Serious[1]	Not serious	Not serious[4]	Serious[3]	66	22	MD -0.63 (-1.37, 0.10)	Low
Number of adverse events: lower values favour intervention									
Shao 2015	RCT	Serious[1]	Not serious	Not serious[4]	Very serious[5]	66	22	RR 1.40 (0.79, 2.47)	Very low

1. High risk of bias to lack of reported blinding
2. I^2 >40%
3. Non-significant result
4. 3 comparisons in one trial
5. 95% CI crosses one line of a defined MID interval

G.7.2.4 Cholinesterase inhibitor withdrawal

Quality assessment						No of patients		Effect estimate	
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Withdrawal	Continuation	Effect size (95% CI)	Quality
Cognition (MMSE): lower values favour continuation									
Hermann 2016; Howard 2012[a]	RCT	Not serious	Not serious	Serious[1]	Serious[2]	73	75	MD -1.84 (-3.74, 0.06)	Low
Activities of daily living (ADCS-ADL/BADLS): higher values favour continuation									
Hermann 2016; Howard 2012[a]	RCT	Not serious	Not serious	Not serious	Serious[3]	74	74	SMD 0.21 (-0.11, 0.54)	Moderate
Behavioural and psychological symptoms (NPI): higher values favour continuation									
Hermann 2016; Howard 2012[a]	RCT	Not serious	Not serious	Serious[1]	Serious[2]	73	75	MD 0.23 (-7.79, 8.26)	Low
Quality of life (DEMQOL): lower values favour continuation									
Howard 2012[a]	RCT	Not serious	Not serious	Not serious	Serious[2]	55	54	MD -0.50 (-5.47, 4.46)	Moderate
GHQ-12: lower values favour continuation									
Howard 2012[a]	RCT	Not serious	Not serious	Not serious	Serious[2]	45	51	MD 0.55 (-0.71, 1.81)	Moderate
Entry to care home: lower numbers favour continuation									
Howard 2012	RCT	Not serious	Not serious	N/A	Serious[2]	76	73	HR 1.22 (0.78, 1.90)	Moderate

1. I^2 >40%
2. Non-significant result
3. 95% CI crosses one line of a defined MID interval

a: extracted from additional data (see appendix E)

G.7.2.5 Cholinesterase inhibitor switch to memantine

Quality assessment						No of patients		Effect estimate	
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Memantine	Continuation	Effect size (95% CI)	Quality
Cognition (MMSE): lower values favour continuation									
Howard 2012[a]	RCT	Not serious	Not serious	Not serious	Serious[2]	51	54	MD -0.47 (-1.77, 0.83)	Moderate
Activities of daily living (ADCS-ADL/BADLS): higher values favour continuation									
Howard 2012[a]	RCT	Not serious	Not serious	Not serious	Serious[2]	51	54	MD 0.21 (-2.91, 3.34)	Moderate
Behavioural and psychological symptoms (NPI): higher values favour continuation									
Howard 2012[a]	RCT	Not serious	Not serious	Serious[1]	Serious[2]	51	54	MD -9.28 (-20.49, 1.93)	Low
Quality of life (DEMQOL): lower values favour continuation									
Howard 2012[a]	RCT	Not serious	Not serious	Not serious	Serious[2]	51	54	MD 2.62 (-3.43, 8.66)	Moderate
GHQ-12: lower values favour continuation									
Howard 2012[a]	RCT	Not serious	Not serious	Serious[1]	Serious[2]	47	51	MD -0.07 (-2.00, 1.86)	Low
Entry to care home: lower numbers favour continuation									
Howard 2012	RCT	Not serious	Not serious	N/A	Serious[2]	76	73	HR 1.40 (0.90, 2.20)	Moderate

1. I^2 >40%
2. Non-significant result

a: extracted from additional data (see appendix E)

G.7.3 Pharmacological management of Parkinson's disease dementia

- What is the comparative effectiveness of donepezil, galantamine, memantine and rivastigmine for cognitive enhancement in dementia associated with Parkinson's disease?

G.7.3.1 Parkinson's disease dementia – cholinesterase inhibitors

PDD – cholinesterase inhibitor vs. placebo: adverse events

Quality assessment						No of patients		Effect		Quality
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	ChI	Placebo	Relative (95% CI)	Absolute (95% CI)	
Any adverse events – cholinesterase inhibitors (probability of experiencing ≥1; follow-up 10 to 24 weeks; lower is better)										
4[1–4]	RCT	not serious	not serious	not serious	serious[5]	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)	⊕⊕⊕O MODERATE
Any adverse events – donepezil (probability of experiencing ≥1; follow-up 10 to 24 weeks; lower is better)										
3[1,2,4]	RCT	not serious	not serious	not serious	serious[5]	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 events – rivastigmine (probability of experiencing ≥1; follow-up 24 weeks; lower is better)										
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 adverse events – cholinesterase inhibitors (probability of experiencing ≥1; follow-up 24 weeks; lower is better)										
2[2,3]	RCT	not serious	serious[6]	not serious	serious[5]	114/739 (15.4%)	48/352 (13.6%)	RR 1.12 (0.72 to 1.73)	18 more per 1000 (from 39 fewer to 100 more)	⊕⊕OO LOW
Serious adverse events – donepezil (probability of experiencing ≥1; follow-up 24 weeks; lower is better)										
1[2]	RCT	not serious	N/A	not serious	serious[5]	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 adverse events – rivastigmine (probability of experiencing ≥1; follow-up 24 weeks; lower is better)										
1[3]	RCT	not serious	N/A	not serious	serious[5]	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 events requiring treatment withdrawal – cholinesterase inhibitors (probability of experiencing; follow-up 24 weeks; lower is better)										
3[1–3]	RCT	not serious	not serious	not serious	serious[5]	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)	⊕⊕⊕O MODERATE
Adverse events requiring treatment withdrawal – donepezil (probability of experiencing; follow-up 24 weeks)										
2[1,2]	RCT	not serious	not serious	not serious	serious[5]	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 events requiring treatment withdrawal – rivastigmine (probability of experiencing; 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

Quality assessment						No of patients		Effect		Quality
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Chl	Placebo	Relative (95% CI)	Absolute (95% CI)	
Hallucinations – cholinesterase inhibitors (probability of experiencing; follow-up 24 weeks; lower is better)										
2[2,3]	RCT	not serious	not serious	not serious	serious[5]	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)	⊕⊕⊕O MODERATE
Hallucinations – donepezil (probability of experiencing; follow-up 24 weeks; lower is better)										
1[2]	RCT	not serious	N/A	not serious	serious[5]	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 – rivastigmine (probability of experiencing; follow-up 24 weeks; lower is better)										
1[3]	RCT	not serious	N/A	not serious	serious[5]	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)	⊕⊕⊕O MODERATE

[1] Aarsland 2002
[2] Dubois 2012; data for 2 active treatment groups were combined (donepezil 5mg and 10mg)
[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

PDD – rivastigmine patches vs. rivastigmine capsules: adverse events

Quality assessment						No of patients		Effect		Quality
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Rivastigmine patches	Rivastigmine capsules	Relative (95% CI)	Absolute (95%CI)	
Any adverse events (probability of experiencing ≥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 adverse events (probability of experiencing ≥1; follow-up 76 weeks; 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 events requiring treatment withdrawal (probability of experiencing; follow-up 76 weeks; 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
Hallucinations (probability of experiencing ; follow-up 76 weeks)										
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

PDD – cholinesterase inhibitor vs. placebo: cognitive function

Quality assessment						No of patients		Effect	Quality
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Chl	Placebo	Mean difference (95% CI)	
MMSE – cholinesterase inhibitors (follow-up 10 to 24 weeks; range of scores: 0-30; higher is better)									
4[1–4]	RCT	not serious	not serious	not serious	not serious	752	367	1.36 higher (0.95 to 1.77 higher)	⊕⊕⊕⊕ HIGH
MMSE – donepezil (follow-up 10 to 24 weeks; range of scores: 0-30; higher is better)									
3[1,2,4]	RCT	not serious	not serious	not serious	not serious	417	201	1.58 higher (1.06 to 2.1 higher)	⊕⊕⊕⊕ HIGH
MMSE – rivastigmine (follow-up 24 weeks; range of scores: 0-30; higher is better)									
1[3]	RCT	not serious	N/A	not serious	not serious	335	166	1 higher (0.33 to 1.67 higher)	⊕⊕⊕⊕ HIGH
ADAS-cog – cholinesterase inhibitors (follow-up 10 to 24 weeks; range of scores: 0-70; lower is better)									
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 – donepezil (follow-up 10 to 24 weeks; range of scores: 0-70; lower is better)									
2[2,4]	RCT	not serious	not serious	not serious	serious[5]	360	185	1.5 lower (3.28 lower to 0.27 higher)	⊕⊕⊕O MODERATE
ADAS-cog – rivastigmine (follow-up 24 weeks; range of scores: 0-70; lower is better)									
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) – cholinesterase inhibitors (follow-up 10 to 24 weeks; range of scores: 0-144; higher is better)[6]									
2[3,4]	RCT	not serious	not serious	not serious	very serious[5,7]	35	31	3.39 higher (4.06 lower to 10.84 higher)	⊕⊕OO LOW
MDRS (total score) – donepezil (follow-up 10 weeks; range of scores: 0-144; higher is better)									
1[4]	RCT	not serious	N/A	not serious	very serious[5,7]	19	19	0.2 lower (11.44 lower to 11.04 higher)	⊕⊕OO LOW
MDRS (total score) – rivastigmine (follow-up 24 weeks; range of scores: 0-144; higher is better)[6]									
1[3]	RCT	serious[7]	N/A	not serious	serious[5]	16	12	6.21 higher (3.75 lower to 16.17 higher)	⊕⊕OO LOW
Clock drawing test – rivastigmine (follow-up 24 weeks; range of scores: 0-10; higher is better)									
1[3]	RCT	serious[7]	N/A	not serious	serious[5]	49	30	1.1 higher (0.01 lower to 2.21 higher)	⊕⊕OO LOW
D-KEFS verbal fluency test (total score) – rivastigmine (follow-up 24 weeks; measured by number of correct responses; 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 fluency test (letter fluency) – donepezil (follow-up 24 weeks; higher is better)									

Quality assessment						No of patients		Effect	Quality
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Chl	Placebo	Mean difference (95% CI)	
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 fluency test (category fluency) – donepezil (follow-up 24 weeks; higher is better)									
1[2]	RCT	not serious	N/A	not serious	not serious	307	152	3.93 higher (2.05 to 5.81 higher)	⊕⊕⊕⊕ HIGH
D-KEFS verbal fluency test (category switching) – donepezil (follow-up 24 weeks; higher is better)									
1[2]	RCT	not serious	N/A	not serious	serious[5]	307	152	1.09 higher (0.79 lower to 2.97 higher)	⊕⊕⊕O MODERATE
CDR – rivastigmine (follow-up 24 weeks; measured with: milliseconds; lower is better)									
1[3]	RCT	not serious	N/A	not serious	serious[5]	328	158	173.7 lower (471.23 lower to 123.83 higher)	⊕⊕⊕O MODERATE
BTA – donepezil (follow-up 24 weeks; range of scores: 0-20; higher is better)									
1[2]	RCT	serious[8]	N/A	not serious	not serious	221	111	0.88 higher (0.4 to 1.37 higher)	⊕⊕⊕O MODERATE

[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

Quality assessment						No of patients		Effect	Quality
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Rivastigmine patches	Rivastigmine capsules	Mean difference (95% CI)	
MDRS (total score) (follow-up 24 weeks; range of scores 0-144; higher is better)									
1[1]	RCT	serious[2]	N/A	not serious	serious[3]	273	273	2.1 lower (4.27 lower to 0.07 higher)	⊕⊕OO LOW
MDRS (total score) (follow-up 76 weeks; range of scores 0-144; higher is better)									
1[1]	RCT	serious[2]	N/A	not serious	not serious	273	273	5.3 lower (8.17 to 2.43 lower)	⊕⊕⊕O MODERATE

[1] Emre 2014
[2] Open-label study
[3] At a 95% confidence level, data are consistent with appreciable harm, appreciable benefit or no difference

PDD – cholinesterase inhibitor vs. placebo: global assessment

Quality assessment						No of patients		Effect (95%CI)	Quality
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	ChI	Placebo		
Global function – cholinesterase inhibitors (follow-up 10 to 24 weeks; measured with: CIBIC+, ADCS-CGIC or CGIC; range of scores: 1-7; lower is better)									
4[1–4]	RCT	not serious	not serious	not serious	serious[5]	707	366	SMD 0.3 lower (0.42 to 0.17 lower)	⊕⊕⊕O MODERATE
Global response – cholinesterase inhibitors (at least minimal improvement; follow-up 10 to 24 weeks; measured with: CIBIC+ or ADCS-CGIC; higher is better)									
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 – donepezil (at least minimal improvement; follow-up 10 to 24 weeks; measured with: CIBIC+; higher is better)									
2[1,2]	RCT	not serious	not serious	not serious	serious[5]	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 – rivastigmine (at least minimal improvement; follow-up 24 weeks; measured with: ADCS-CGIC; higher is better)									
1[3]	RCT	not serious	N/A	not serious	serious[5]	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)	⊕⊕⊕O MODERATE
CIBIC+ – donepezil (follow-up 10 to 24 weeks; range of scores: 1-7; lower is better)									
2[1,2]	RCT	not serious	serious[6]	not serious	serious[5]	359	182	MD 0.43 lower (0.93 lower to 0.08 higher)	⊕⊕OO LOW
CGIC – donepezil (follow-up 10 weeks; range of scores: 1-7; lower is better)									
1[4]	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 score) – donepezil (follow-up 10 weeks; range of scores: 0-199; lower is better)									
1[4]	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 – rivastigmine (follow-up 24 weeks; range of scores: 1-7; lower is better)									
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] 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
[8]CI cross MID of 7.3 points (Schrag et al., 2006)

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

Quality assessment	No of patients	Effect (95% CI)	Quality

No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Chl	Placebo		
ADL – cholinesterase inhibitors (follow-up 24 weeks; measured with: ADCS-ADL or DAD; higher is better)									
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 – rivastigmine (follow-up 24 weeks; range of scores: 0-78; higher is better)									
1[2]	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 for 2 active treatment groups were combined (donepezil 5mg and 10mg). Mean and standard deviation calculated from data reported in paper
[2] Emre 2004
[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 assessment						No of patients		Effect	Quality
No of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Rivastigmine patches**	**Rivastigmine capsules**	**Mean difference (95% CI)**	
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 (follow-up 76 weeks; range of scores: 0-78; higher is better)									
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 study
[3] At a 95% confidence level, data are consistent with appreciable harm, appreciable benefit or no difference

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)**	
NPI-10 item – cholinesterase inhibitors (follow-up 24 weeks; range of scores: 0-120; lower is better)									
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
NPI-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 – rivastigmine (follow-up 24 weeks; range of scores: 0-120; 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 – donepezil (follow-up 10 weeks; lower is better)									
2[5,6]	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; data for 2 active treatment groups were combined (donepezil 5mg and 10mg). Mean and standard deviation calculated from data reported in paper
[2] Emre 2004
[3] 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
[4] At a 95% confidence level, data are consistent with appreciable harm, appreciable benefit or no difference
[5] Aarsland 2002
[6] Ravina 2005
[7]Data for this outcome not reported in 2 large RCTs (Dubois 2012 and Emre 2004). Papers stated no significant difference between groups
[8]CI 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

Quality assessment						No of patients		Effect	Quality
No of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Rivastigmine patches**	**Rivastigmine capsules**	**Mean difference (95% CI)**	
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 (follow-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] Emre 2014
[2] Open-label study
[3] At a 95% confidence level, data are consistent with appreciable harm, appreciable benefit or no difference
[4]CI do not cross MID between 3.25 (Horvath et al., 2015) and 5 points (Schrag et al., 2006)

G.7.3.2 Parkinsons disease dementia – memantine

PDD – memantine vs. placebo: adverse events

Quality assessment						No of patients		Effect		Quality
No of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Memantine**	**Placebo**	**Relative (95% CI)**	**Absolute (95% CI)**	

Any adverse events (probability of experiencing ≥1; follow-up 16 to 24 weeks, lower is better)										
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)										
2[1,2]	RCT	not serious	not serious	not serious	very serious[3,4]	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 events requiring treatment withdrawal (probability of experiencing; follow-up 24 weeks, lower is better)										
1[1]	RCT	not serious	N/A	not serious	very serious[3,4]	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

[1] Emre 2010; data reported for PDD population only; study also included people with 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] Very small numbers of events

PDD – memantine vs. placebo: cognitive function

Quality assessment						No of patients		Effect	Quality
No of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Memantine**	**Placebo**	**Mean difference (95% CI)**	
MMSE (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 test (follow-up 24 weeks; range of scores: 0-10; higher is better)									
1[4]	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; data reported for end of drug treatment phase (16 weeks)
[2] At a 95% confidence level, data are consistent with appreciable benefit, appreciable harm or no difference
[3] Very small numbers of participants in the study
[4] Emre 2010; data reported for PDD population only; study also included people with DLB

PDD – memantine vs. placebo: global assessment

Quality assessment						No of patients		Effect (95% CI)	Quality
No of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Memantine**	**Placebo**		
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; follow-up 16 weeks; higher is better)									
1[3]	RCT	not serious	N/A	not serious	very serious[2,4]	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; 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

Quality assessment						No of patients		Effect	Quality
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Memantine	Placebo	Mean difference (95% CI)	
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; 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

PDD – memantine vs. placebo: carer-reported outcomes

Quality assessment						No of patients		Effect	Quality
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Memantine	Placebo	Mean difference (95% CI)	
ZBI (follow-up 16 to 24 weeks; lower is better)[1]									
2[2,3]	RCT	not serious	not serious	not serious	serious[4]	71	70	3.4 lower (7.21 lower to 0.42 higher)	⊕⊕⊕O MODERATE

[1] Data from Leroi 2009 reported in a secondary publication (Leroi 2014)
[2] Leroi 2009; data reported for end of drug treatment phase (16 weeks)
[3] Emre 2010; data reported for PDD population only; study also included people with DLB
[4] At a 95% confidence level, data are consistent with appreciable benefit, appreciable harm or no difference

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

Quality assessment						No of patients		Effect	Quality
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Memantine	Placebo	Mean difference (95% CI)	
NPI 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 (follow-up 16 weeks; range of scores: 0-120; lower is better)									
1[2]	RCT	not serious	N/A	not serious	very serious[3,4]	10	14	MD 2.00 lower (11.64 lower to 7.64 higher)	⊕⊕OO LOW
UPDRS III (follow-up 16 to 24 weeks; lower is better)									

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; data reported for PDD population only; study also included people with DLB
[2] Leroi 2009; data reported for end of drug treatment phase (16 weeks)
[3] At a 95% confidence level, data are consistent with appreciable benefit, appreciable harm or no difference
[4] Data from a single very small study
[5]CI cross MID between 3.25 (Horvath et al 2015) and 5 points (Schrag et al., 2006)

G.7.3.3 Dementia with Lewy bodies – cholinesterase inhibitors

DLB – cholinesterase inhibitor vs. placebo: adverse events

Quality assessment						No of patients		Effect		Quality
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	ChI	Placebo	Relative (95% CI)	Absolute (95% CI)	
Any adverse events – cholinesterase inhibitors (probability of experiencing ≥1; follow-up 12 to 20 weeks)										
3[1–3]	RCT	not serious	not serious	not serious	serious[4]	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 events – donepezil (probability of experiencing ≥1; follow-up 12 weeks)										
2[1,2]	RCT	not serious	not serious	not serious	serious[4]	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 events – rivastigmine (probability of experiencing ≥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 adverse events – cholinesterase inhibitors (probability of experiencing ≥1; follow-up 12 to 20 weeks)										
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 adverse events – donepezil (probability of experiencing ≥1; follow-up 12 weeks)										
2[1,2]	RCT	not serious	not serious	not serious	serious[4]	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 adverse events – rivastigmine (probability of experiencing ≥1; follow-up 20 weeks)										
1[3]	RCT	not serious	N/A	not serious	serious[4]	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 requiring treatment withdrawal – cholinesterase inhibitors (probability of experiencing; follow-up 12 to 20 weeks)										
3[1–3]	RCT	not serious	not serious	not serious	serious[4]	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 requiring treatment withdrawal – donepezil (probability of experiencing; follow-up 12 weeks)										
2[1,2]	RCT	not serious	not serious	not serious	serious[4]	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

Quality assessment						No of patients		Effect		Quality
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Chl	Placebo	Relative (95% CI)	Absolute (95% CI)	
Adverse events requiring treatment withdrawal – rivastigmine (probability of experiencing; 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; data for 2 active treatment groups were combined (donepezil 5mg and 10mg)
[2] Mori 2012; data for 3 active treatment groups were combined (donepezil 3mg, 5mg and 10mg)
[3] McKeith 2000
[4] At a 95% confidence level, data are consistent with appreciable harm, appreciable benefit or no difference

DLB – cholinesterase inhibitor vs. placebo: cognitive function

Quality assessment						No of patients		Effect	Quality
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Chl	Placebo	Mean difference (95% CI)	
MMSE – cholinesterase inhibitors (follow-up 12 to 20 weeks; range of scores: 0-30; higher is better)									
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 – donepezil (follow-up 12 weeks; range of scores: 0-30; higher 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 – rivastigmine (follow-up 20 weeks; range of scores: 0-30; higher is better)									
1[2]	RCT	not serious	N/A	not serious	serious[5]	59	61	1.24 higher (0.28 lower to 2.76 higher)	⊕⊕⊕O MODERATE

[1] Ikeda 2015; data for 2 active treatment groups were combined (donepezil 5mg and 10mg)
[2] McKeith 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] i^2 >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: global assessment

Quality assessment						No of patients		Effect (95% CI)	Quality
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Chl	Placebo		
CIBIC+ – donepezil (follow-up 12 weeks; range of scores: 1-7; lower is better)[1]									
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+ – donepezil (at least minimal improvement; 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 calculated from data presented in paper
[2] Mori 2012; data for 3 active treatment groups were combined (donepezil 3mg, 5mg and 10mg)

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

Quality assessment						No of patients		Effect	Quality
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	ChI	Placebo	Mean difference (95% CI)	
ZBI - donepezil (follow-up 12 weeks; lower is better)									
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 for 2 active treatment groups were combined (donepezil 5mg and 10mg)
[2] Mori 2012; data for 3 active treatment groups were combined (donepezil 3mg, 5mg and 10mg)

DLB – 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	ChI	placebo	Mean difference (95% CI)	
NPI-10 item – cholinesterase inhibitors (follow-up 12 to 20 weeks; range of scores: 0-120; lower is better)[1]									
3[2–4]	RCT	not serious	serious[5]	not serious	serious[6]	243	129	2.06 lower (7.15 lower to 3.02 higher)	⊕⊕OO LOW
NPI-10 item – donepezil (follow-up 12 weeks; range of scores: 0-120; lower is better)[1]									
2[2,4]	RCT	not serious	serious[5]	not serious	serious[6]	196	76	1.54 lower (9.37 lower to 6.29 higher)	⊕⊕OO LOW
NPI-10 item – rivastigmine (follow-up 20 weeks; range of scores: 0-120; lower 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 – cholinesterase inhibitors (follow-up 12 to 20 weeks; range of scores: 0-48; lower is better)[7]									
2[3,4]	RCT	not serious	not serious	not serious	not serious	161	93	2.49 lower (4.64 to 0.33 lower)	⊕⊕⊕⊕ HIGH
NPI-4 item – donepezil (follow-up 12 weeks; range of scores: 0-48; lower is better)[7]									
1[4]	RCT	not serious	N/A	not serious	not serious	102	32	3.59 lower (6.93 to 0.25 lower)	⊕⊕⊕⊕ HIGH
NPI-4 item – rivastigmine (follow-up 20 weeks; range of scores: 0-48; lower 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 – donepezil (follow-up 12 weeks; range of scores: 0-24; lower is better)[8]									
2[2,4]	RCT	not serious	serious[5]	not serious	serious[6]	196	76	2.3 lower (6.32 lower to 1.72 higher)	⊕⊕OO LOW

UPDRS III – cholinesterase inhibitors (follow-up 12 weeks; lower is better)[1]									
2[2,4]	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 – donepezil (follow-up 12 weeks; lower is 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 for this outcome in Ikeda 2015; calculated from SE reported in paper
[2] Ikeda 2015; data for 2 active treatment groups were combined (donepezil 5mg and 10mg)
[3] McKeith 2000
[4] Mori 2012; data for 3 active treatment groups were combined (donepezil 3mg, 5mg and 10mg)
[5] i^2 >40% between studies
[6] At a 95% confidence level, data are consistent with appreciable benefit, appreciable harm or no difference
[7] NPI 4-item consists of 4 NPI domains – hallucinations, delusions, dysphoria and apathy
[8] NPI 2-item consists of 2 NPI domains – hallucinations and cognitive fluctuation
[9] Data for outcome not presented in McKeith 2000. Study reported no significant difference between groups
[10] CI do not cross MID between 3.25 (Horvath et al., 2015) and 5 points (Schrag et al., 2006)

G.7.3.4 Dementia with Lewy bodies – memantine

DLB – memantine vs. placebo: adverse events

Quality assessment						No of patients		Effect		Quality
No of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Memantine**	**Placebo**	**Relative (95% CI)**	**Absolute (95% CI)**	
Any adverse events (probability of experiencing ≥1; follow-up 24 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 adverse events (probability of experiencing ≥1; follow-up 24 weeks)										
1[1]	RCT	not serious	N/A	not serious	very serious[2,3]	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 events requiring treatment withdrawal (probability of experiencing; follow-up 24 weeks)										
1[1]	RCT	not serious	N/A	not serious	very serious[2,3]	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

Quality assessment	No of patients	Effect	Quality

No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Memantine	Placebo	Mean difference (95% CI)	
Clock drawing test (follow-up 24 weeks; range of scores: 0-10; higher is better)									
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; data reported for DLB population only; study also included people with PDD
[2] At a 95% confidence level, data are consistent with appreciable benefit, appreciable harm or no difference

DLB – memantine vs. placebo: global assessment

Quality assessment						No of patients		Effect	Quality
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Memantine	Placebo	Mean difference (95% CI)	
ADCS-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; data reported for DLB population only; study also included people 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

Quality assessment						No of patients		Effect	Quality
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Memantine	Placebo	Mean difference (95% CI)	
ADCS-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; data reported for DLB population only; study also included people with PDD
[2] Wide 95% confidence intervals, data are consistent with appreciable benefit, appreciable harm or no difference

DLB – memantine vs. placebo: carer-reported outcomes

Quality assessment						No of patients		Effect	Quality
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Memantine	Placebo	Mean difference (95% CI)	
ZBI (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; data reported for DLB population only; study also included people 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

Quality assessment						No of patients		Effect	Quality
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Memantine	Placebo	Mean difference (95% CI)	
NPI-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 weeks; lower is better)									
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; data reported for DLB population only; study also included people with PDD
[2] Wide 95% confidence intervals, data are consistent with appreciable benefit, appreciable harm or no difference
[3]CI cross the MID between 3.25 (Horvath et al., 2015) and 5 points (Schrag et al., 2006)

G.7.3.5 Mixed population (PDD or DLB) – cholinesterase inhibitors

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

Quality assessment						No of patients		Effect		Quality
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	ChI	Placebo	Relative (95% CI)	Absolute (95% CI)	
Any adverse events – cholinesterase inhibitors (probability of experiencing ≥1; follow-up 10 to 24 weeks; lower is better)										
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 events – donepezil (probability of experiencing ≥1; follow-up 10 to 24 weeks; 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 events – rivastigmine (probability of experiencing ≥1; follow-up 20 to 24 weeks; lower 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 adverse events – cholinesterase inhibitors (probability of experiencing ≥1; follow-up 12 to 24 weeks; lower is better)										
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 adverse events – donepezil (probability of experiencing ≥1; follow-up 12 to 24 weeks; lower is better)										
3[2,4,6]	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 adverse events – rivastigmine (probability of experiencing ≥1; follow-up 20 to 24 weeks; lower 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 requiring treatment withdrawal – cholinesterase inhibitors (probability of experiencing; follow-up 10 to 24 weeks; lower is better)										

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 events requiring treatment withdrawal – donepezil (probability of experiencing; follow-up 10 to 24 weeks; lower is better)										
4[1,2,4,6]	RCT	not serious	not serious	not serious	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 events requiring treatment withdrawal – rivastigmine (probability of experiencing; follow-up 20 to 24 weeks; lower is better)										
2[3,5]	RCT	not serious	not serious	not serious	not serious	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 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
[8] At a 95% confidence level, data are consistent with appreciable benefit, appreciable harm or no difference

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

Quality assessment						No of patients		Effect	Quality
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	ChI	Placebo	Mean difference (95% CI)	
MMSE – cholinesterase inhibitors (follow-up 10 to 24 weeks; range of scores: 0-30; higher is better)									
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; range of scores: 0-30; higher is better)									
5[1,2,4,6,7]	RCT	not serious	not serious	not serious	not serious	614	276	1.68 higher (1.24 to 2.11 higher)	⊕⊕⊕⊕ HIGH
MMSE – rivastigmine (follow-up 20 to 24 weeks; range of scores: 0-30; higher is better)									
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

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

Quality assessment						No of patients		Effect (95% CI)	Quality
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Chl	Placebo		
Global function – cholinesterase inhibitors (follow-up 10 to 24 weeks; measured with: CIBIC+, ADCS-CGIC or CGIC; range of scores: 1-7; lower is better)									
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 – donepezil (follow-up 10 to 24 weeks; measured with: CIBIC+, ADCS-CGIC or CGIC; range of scores: 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 – cholinesterase inhibitors (at least minimal improvement; follow-up 10 to 24 weeks; measured with: CIBIC+ or ADCS-CGIC; higher is better)									
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 – donepezil (at least minimal improvement; follow-up 10 to 24 weeks; measured with: CIBIC+ or ADCS-CGIC; higher is better)									
3[1,2,4]	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; 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] *Mori 2012; data for 3 active treatment groups were combined (donepezil 3mg, 5mg and 10mg)*
[5] *Ravina 2005*
[6] *Heterogeneity >40% between studies*

PDD/DLB – 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)	
NPI-10 item – cholinesterase inhibitors (follow-up 12 to 24 weeks; range of scores: 0-120; lower is better)[1]									
5[2–6]	RCT	not serious[7]	not serious	not serious	not serious	931	465	1.49 lower (2.69 to 0.29 lower)	⊕⊕⊕⊕ HIGH
NPI-10 item – donepezil (follow-up 12 to 24 weeks; range of scores: 0-120; lower is better)[1]									
3[2,4,6]	RCT	not serious[7]	serious[8]	not serious	serious[9]	550	246	0.92 lower (2.54 lower to 0.69 higher)	⊕⊕OO LOW
NPI-10 item – rivastigmine (follow-up 20 to 24 weeks; range of scores: 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 – donepezil (follow-up 24 weeks; lower is better)									
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

[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
[12]Data for outcome not reported in 3 large RCTs (Dubois 2012, Emre 2004 and McKeith 2000). Papers stated no significant difference between groups
[13]CI do not cross the MID between 3.25 (Horvath et al., 2015) and 5 points (Schrag et al., 2006)

G.7.3.6 Mixed population (PDD or DLB) – memantine

PDD/DLB – memantine vs. placebo: adverse events

Quality assessment						No of patients		Effect		Quality
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Memantine	Placebo	Relative (95% CI)	Absolute (95% CI)	
Any adverse events (probability of experiencing ≥1; follow-up 16 to 24 weeks; lower is better)										
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 adverse events (probability of experiencing ≥1; follow-up 16 to 24 weeks; lower is better)										
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 events requiring treatment withdrawal (probability of experiencing; follow-up 16 to 24 weeks; lower is better)										
2[2,4]	RCT	not serious	not serious	serious[5]	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 reported for total population (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: cognitive outcomes

Quality assessment						No of patients		Effect	Quality
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Memantine	Placebo	Mean difference (95% CI)	
MMSE (follow-up 16 to 24 weeks; range of scores: 0-30; higher is better)									
2[1,2]	RCT	not serious	not serious	serious[3]	serious[3]	40	47	1.56 higher (0.17 lower to 3.28 higher)	⊕⊕OO LOW

[1] Aarsland 2009
[2] Leroi 2009; data reported for end of drug treatment phase (16 weeks)
[3] Both studies included people who were also taking a cholinesterase inhibitor
[4] At a 95% confidence level, data are consistent with appreciable harm, appreciable benefit or no difference

PDD/DLB – memantine vs. placebo: global assessment

Quality assessment						No of patients		Effect	Quality
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Memantine	Placebo	Standardised mean difference (95% CI)	
Global function (follow-up 24 weeks; measured with: ADCS-CGIC or CGIC; range of scores: 1-7; lower is better)									
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)

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

Quality assessment						No of patients		Effect	Quality
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Memantine	Placebo	Standardised mean difference (95% CI)	
ADL (follow-up 24 weeks; measured with: ADCS-ADL or DAD; higher is better)									
2[1,2]	RCT	not serious	not serious	not serious	serious[3]	123	130	0.13 higher (0.12 lower to 0.38 higher)	⊕⊕⊕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

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

Quality assessment						No of patients		Effect	Quality
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Memantine	Placebo	Mean difference (95% CI)	
ZBI (follow-up 16 to 24 weeks; lower is better)									
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; data reported for total population (PDD and DLB)*
[2] *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: other non-cognitive outcomes

Quality assessment						No of patients		Effect (95% CI)	Quality
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Memantine	Placebo		
NPI (follow-up 16 to 24 weeks; measured with: NPI-10 item or NPI 12-item; lower is better)[1]									
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; lower is better)									
2[2,3]	RCT	not serious	not serious	not serious	not serious[5]	131	141	MD 0.28 higher (1.28 lower to 1.85 higher)	⊕⊕⊕⊕ HIGH

[1] *Data from Leroi 2009 could not be included in this analysis due to inconsistent outcome reporting*
[2] *Aarsland 2009*
[3] *Emre 2010; data reported for total population (PDD and DLB)*
[4] *At a 95% confidence level, data are consistent with appreciable harm, appreciable benefit or no difference*
[5] *CI do not cross the MID between 3 (Horvath et al., 2015) and 5 points (Schrag et al., 2006)*

Network meta-analyses

Any adverse events

Quality assessment					Quality
Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	
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, interventions, comparator and outcomes are as defined in protocol

Serious adverse events

Quality assessment					Quality
Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	
Serious adverse events					

Quality assessment					
Number of RCTs	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Quality**
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, interventions, 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 withdrawal					
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, interventions, 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, McKeith 2000, Aarsland 2009, Emre 2010	Not serious	Not serious	Not serious[1]	Not serious	High

1. Considered not serious as population, interventions, 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)					

Quality assessment					
Number of RCTs	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Quality**
7 Aarsland 2002, Dubois 2012, Mori 2012, Ravina 2005, Emre 2004, Aarsland 2009, Emre 2010	Not serious	Serious[1]	Not serious[2]	Not serious	Moderate

1. Considerable between study heterogeneity (i^2>40%)
2. Considered not serious as population, interventions, comparator and outcomes are as defined in protocol

NPI

Quality assessment					
Number of RCTs	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Quality**
Change in NPI scores					
8 Dubois 2012, Ikeda 2015, Mori 2012, Emre 2004, McKeith 2000, Aarsland 2009, Emre 2010, Leroi 2009	Not serious	Not serious	Not serious[1]	Not serious	High

1. 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

G.7.4 Cholinesterase inhibitors and memantine for types of dementia other than typical Alzheimer's disease

- How effective are cholinesterase inhibitors and memantine for types of dementia other than typical Alzheimer's disease?

G.7.4.1 Vascular dementia

Cholinesterase inhibitors versus placebo

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	AChEI	Placebo	Summary of results	
Cognitive outcomes – global cognition									
MMSE (higher values = better score)									
4 (Ballard 2008, Black 2003, Mok 2007, Roman 2010)	RCT	Not serious	Not serious	Not serious	Not serious	1,417	884	MD 0.58 (0.30, 0.86)	High
ADAS-cog (lower values = better score)									
4 (Ballard 2008, Black 2003, Roman 2010, Wilkinson 2003)	RCT	Not serious	Not serious	Serious[1]	Not serious	1,719	1,015	MD -1.36 (-2.03, -0.70)	Moderate
ADAS-cog-11 (lower values = better score)									
2 (Auchus 2007, Small 2003)	RCT	Not serious	Not serious	Not serious	Not serious	486	440	MD -1.59 (-2.39, -0.78)	High
Vascular Dementia Assessment Scale – cognitive subscale (lower values = better score)									
1 (Roman 2010)	RCT	Not serious	Not serious	N/A	Not serious	535	283	MD -1.15 (-1.99, -0.31)	High
EXIT-25 (lower values = better score)									
2 (Auchus 2007, Roman 2010)	RCT	Not serious	Not serious	Serious[1]	Serious[2]	991	692	MD -0.57 (-1.40, 0.25)	Low
Neuropsychiatric symptoms									
NPI (lower values = better score)									
2 (Auchus 2007, Mok 2007)	RCT	Not serious	Not serious	Not serious	Not serious	376	381	MD 1.76 (0.28, 3.24)	High
NPI-12 (lower values = better score)									

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	AChEI	Placebo	Summary of results	
1 (Ballard 2008)	RCT	Not serious	Not serious	N/A	Serious[2]	364	342	MD 0.40 (-1.36, 2.16)	Moderate
Global assessment									
Clinician's Global Impression of Change (lower values = better score)									
1 (Ballard 2008)	RCT	Not serious	Not serious	N/A	Not serious	329	320	MD -0.10 (-3.68, -3.48)	High
Vascular Dementia Assessment Scale (lower values = better score)									
1 (Ballard 2008)	RCT	Not serious	Not serious	N/A	Serious[2]	355	327	MD -1.03 (-2.62, 0.02)	Moderate
Global deterioration scale									
1 (Ballard 2008)	RCT	Not serious	Not serious	N/A	Serious[2]	365	345	MD -0.10 (-2.25, 2.05)	Moderate
Clinical Dementia Rating Sum of Boxes (lower values = better score)									
4 (Black 2003, Mok 2007, Roman 2010, Wilkinson 2003)	RCT	Serious[3]	Not serious	Not serious	Not serious	1,379	696	MD -0.17 (-0.33, -0.00)	Moderate
Functional ability									
ADCS-ADL (higher values = better score)									
2 (Auchus 2007, Ballard 2008)	RCT	Not serious	Not serious	Not serious	Serious[2]	728	716	MD -0.13 (-1.16, 0.90)	Moderate
Instrumental Activities of Daily Living (lower values = better score)									
3 (Black 2003, Mok 2007, Wilkinson 2003)	RCT	Very serious[4]	Not serious	Serious[1]	Serious[2]	751	375	MD -0.38 (-1.04, 0.27)	Very low
Alzheimer's Disease Functional Assessment and Change Scale (lower values = better score)									
2 (Black 2003, Wilkinson 2003)	RCT	Not serious	Not serious	Not serious	Not serious	570	356	MD -0.95 (-1.73, -0.18)	High
Functional Assessment Battery (higher values = better score)									
1 (Mok 2007)	RCT	Not serious	Not serious	N/A	Very serious[5]	20	19	MD -0.40 (-2.13, 1.33)	Low
Disability assessment for Dementia									
1 (Roman 2010)	RCT	Not serious	Not serious	N/A	Serious[2]	628	321	MD 1.77 (-0.10, 3.64)	Moderate
Adverse events									
Any adverse events (lower values = better score)									
5 (Auchus 2007, Black 2003, Mok 2007, Roman 2010, Wilkinson 2003)	RCT	Not serious	Not serious	Not serious	Not serious	1592/1891	884/1128	RR 1.05 (1.01, 1.09)	High
Serious adverse events (lower values = better score									

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	AChEI	Placebo	Summary of results	
5 (Auchus 2007, Ballard 2008, Black 2003, Roman 2010, Wilkinson 2003)	RCT	Not serious	Not serious	Not serious	Serious[6]	337/2019	220/1452	RR 1.11 (0.95, 1.30)	Moderate
Discontinuation due to adverse events (lower values = better score)									
3 (Auchus 2007, Ballard 2008, Mok 2007)	RCT	Not serious	Not serious	Not serious	Not serious	76/779	31/754	RR 2.40 (1.61, 3.59)	High
Mortality (lower values = better scores)									
6 (Auchus 2007, Ballard 2008, Black 2003, Mok 2007, Roman 2010, Wilkinson 2003)	RCT	Not serious	Not serious	Serious[1]	Serious[2]	37/2254	24/1472	RR 0.99 (0.43, 2.30)	Low

1. i^2>40%.
2. Non-significant result.
3. Primary outcomes in some studies presented without measures of dispersion; unclear reporting of sample size in secondary outcomes at endpoint
4. Primary outcomes in some studies only presented in graphs
5. Small sample size and non-significant result.
6. 95% CI crosses one line of a defined MID interval

Memantine versus placebo

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Memantine	Placebo	Summary of results	
Cognitive outcomes - global cognition									
MMSE (higher values = better score)									
1 (Orgogozo 2002)	RCT	Not serious	Not serious	N/A	Not serious	105	108	MD 1.23 (0.23, 2.23)	High
ADAS-cog (lower values = better score)									
2 (Orgogozo 2002, Wilcock 2002[2])	RCT	Not serious	Not serious	Not serious	Not serious	377	375	MD -2.19 (-3.16, -1.21)	High

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Memantine	Placebo	Summary of results	
Behavioural symptoms									
Nurses' Observation Scale for Geriatric Patients (lower values = better score)									
2 (Orgogozo 2002, Wilcock 2002)	RCT	Not serious	Not serious	Not serious	Serious[1]	275	250	MD -0.92 (-2.90, 1.05)	Moderate
Global assessment									
Gottfries-Bråne-Steen scale (lower values = better score)									
2 (Orgogozo 2002, Wilcock 2002)	RCT	Not serious	Not serious	Not serious	Serious[1]	311	284	MD -1.83 (-4.22, 0.56)	Moderate
Clinician's Interview based Impression of Change (lower values = better score)									
1 (Orgogozo 2002)	RCT	Not serious	Not serious	N/A	Serious[1]	114	114	MD -0.29 (-0.66, 0.08)	Moderate
Adverse events									
Any adverse events (lower values = better score)									
1 (Wilcock 2002)	RCT	Not serious	Not serious	N/A	Not serious	226/295	212/284	RR 1.03 (0.94, 1.13)	High
Serious adverse events (lower values = better score)									
1 (Orgogozo 2002)	RCT	Not serious	Not serious	Not serious	Very serious[3]	38/93	40/95	RR 0.97 (0.69, 1.36)	Low

1. Non-significant result.
2. Corrected an error in published results
3. 95% CI crosses two lines of a defined MID interval

Network meta-analyses

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Active	Placebo	Summary of results	
Cognitive outcomes – global cognition									
MMSE (higher values = better score)									

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Active	Placebo	Summary of results	
5 (Ballard 2008, Black 2003, Mok 2007, Orgogozo 2002, Roman 2010)	RCT	Not serious	Not serious	Not serious	Not serious	1,522	992	See appendix H	High
ADAS-cog (lower values = better score)									
6 (Ballard 2008, Black 2003, Orgogozo 2002, Roman 2010, Wilcock 2002, Wilkinson 2003)	RCT	Not serious	Not serious	Serious[1]	Not serious	2,096	1,390	See appendix H	Moderate
Adverse events									
Any adverse events (lower values = better score)									
6 (Auchus 2007, Black 2003, Mok 2007, Roman 2010, Wilcock 2002, Wilkinson 2003)	RCT	Not serious	Not serious	Not serious	Not serious	2,186	1,412	See appendix H	High
Serious adverse events (lower values = better score)									
5 (Auchus 2007, Ballard 2008, Black 2003, Orgogozo 2002, Roman 2010, Wilkinson 2003)	RCT	Not serious	Not serious	Not serious	Serious[2]	2,112	1,547	See appendix H	Moderate
[1i]2>40%. [2]Analysis could not differentiate any treatment groups.									

G.7.4.2 Behavioural variant frontotemporal dementia

Cholinesterase inhibitors versus placebo

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	AChEI	Placebo	Summary of results	
Cognitive outcomes – global cognition									
MMSE (higher values = better score)									
1 (Kerstesz 2008)	RCT	Not serious	Not serious	Not serious	Very serious[1]	17	17	MD 4.40 (-1.02, 9.82)	Low
Dementia Rating Scale (higher values = better score)									
1 (Kerstesz 2008)	RCT	Not serious	Not serious	Not serious	Very serious[1]	17	17	MD 22.00 (-3.37,47.37)	Low

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	AChEI	Placebo	Summary of results	
Neuropsychiatric symptoms									
NPI (lower values = better score)									
1 (Kerstesz 2008)	RCT	Not serious	Not serious	Not serious	Very serious[1]	17	17	MD 5.80 (-7.25, 18.85)	Low
Functional ability									
Functional Assessment Battery (higher values = better score)									
1 (Kerstesz 2008)	RCT	Not serious	Not serious	Not serious	Very serious[1]	17	17	MD 2.50 (-0.99, 5.99)	Low
ADCS-ADL (higher value = better score)									
1 (Kerstesz 2008)	RCT	Not serious	Not serious	Not serious	Very serious[1]	17	17	MD 7.00 (-7.55, 21.55)	Low
Adverse events									
Any adverse events (lower values = better score)									
1 (Kerstesz 2008)	RCT	Not serious	Not serious	Not serious	Very serious[1]	4/18	5/18	RR 0.80 (0.26, 2.50)	Low
Discontinuation due to adverse events (lower values = better score)									
1 (Kerstesz 2008)	RCT	Not serious	Not serious	Not serious	Very serious[1]	1/18	1/18	RR 1.00 (0.07, 14.79)	Low

1. Small sample size and non-significant result.
2. 95% CI crosses two lines of a defined MID interval

Memantine versus placebo

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Memantine	Placebo	Summary of results	
Cognitive outcomes – global cognition									
MMSE (higher values = better score)									
2 (Boxer 2013, Vercelletto 2011)	RCT	Not serious	Not serious	Not serious	Serious[1]	50	55	MD 0.26 (-1.43, 1.95)	Moderate
Mattis Dementia Rating Scale (lower values = better score)									
1 (Vercelletto 2011)	RCT	Not serious	Not serious	N/A	Very serious[2]	18	23	MD 6.30 (-9.55, 22.15)	Low
EXIT-25 (lower values = better score)									
1 (Boxer 2013)	RCT	Not serious	Not serious	N/A	Serious[1]	31	33	MD 1.20 (-1.86, 4.26)	Moderate

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Memantine	Placebo	Summary of results	
Neuropsychiatric symptoms									
NPI (lower values = better score)									
2 (Boxer 2013, Vercelletto 2011)	RCT	Not serious	Not serious	Not serious	Serious[1]	48	55	MD -3.61 (-8.79, 1.57)	Moderate
Global assessment									
Clinician's Interview based Impression of Change (lower values = better score)									
1 (Vercelletto 2011)	RCT	Not serious	Not serious	N/A	Very serious[2]	18	23	MD -0.80 (-1.82, 0.22)	Low
Clinician's Global Impression of Change (lower values = better score)									
1 (Boxer 2013)	RCT	Not serious	Not serious	N/A	Serious[1]	31	33	MD -0.50 (-1.35, 0.35)	Moderate
Clinical Dementia Rating Sum of Boxes (lower values = better score)									
1 (Boxer 2013)	RCT	Not serious	Not serious	N/A	Serious[1]	31	33	MD -0.10 (-2.22, 2.02)	Moderate
Motor function									
Unified Parkinson's disease rating scale (lower values = better score)									
1 (Boxer 2013)	RCT	Not serious	Not serious	N/A	Serious[1]	31	33	MD -0.30 (-3.46, 2.86)	Moderate
Carer burden									
ZBI (lower values = better score)									
1 (Vercelletto 2011)	RCT	Not serious	Not serious	N/A	Very serious[2]	16	23	MD -5.40 (-14.52, 3.72)	Low
Adverse events									
Any adverse events (lower values = better score)									
1 (Vercelletto 2011)	RCT	Not serious	Not serious	N/A	Very serious[4]	8/23	10/26	RR 0.90 (0.43, 1.90)	Low
Serious adverse events (lower values = better score)									
2 (Boxer 2013, Vercelletto 2011)	RCT	Not serious	Not serious	Not serious	Very serious[4]	7/54	12/59	RR 0.65 (0.29,1.48)	Very low
Discontinuation due to adverse events (lower values = better score)									
1 (Vercelletto 2011)	RCT	Not serious	Not serious	N/A	Very serious[4]	3/23	3/26	RR 1.13 (0.25, 5.06)	Low
Mortality (lower values = better scores)									
1 (Vercelletto 2011)	RCT	Not serious	Not serious	N/A	Very serious[2]	2/23	0/26	RR 5.63 (0.28, 111.43)	Low

1. Non-significant result
2. Small sample size and non-significant result.
3. i^2>40%.
4. 95% CI crosses two lines of a defined MID interval

Network meta-analyses

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Active	Placebo	Summary of results	
Cognitive outcomes – global cognition									
MMSE (higher values = better score)									
3 (Boxer 2013, Kertesz 2008, Vercelletto 2011)	RCT	Not serious	Not serious	Not serious	Serious[1]	67	72	See appendix H	Moderate
Neuropsychiatric symptoms									
NPI (lower values = better score)									
3 (Boxer 2013, Kertesz 2008, Vercelletto 2011)	RCT	Not serious	Not serious	Not serious	Serious[1]	65	72	See appendix H	Moderate
Adverse events									
Any adverse events (lower values = better score)									
2 (Kertesz 2008, Vercelletto 2011)	RCT	Not serious	Not serious	N/A	Serious[1]	41	44	See appendix H	Moderate
Discontinuation due to adverse events (lower values = better score)									
2 (Kertesz 2008, Vercelletto 2011)	RCT	Not serious	Not serious	N/A	Serious[1]	41	44	See appendix H	Moderate

1. Analysis could not differentiate any treatment groups.
2. i^2>40%.

G.7.4.3 Semantic variant frontotemporal dementia

Memantine versus placebo

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Memantine	Placebo	Summary of results	
Cognitive outcomes – global cognition									
MMSE (higher values = better score)									
1 (Boxer 2013)	RCT	Not serious	Not serious	N/A	Very serious[1]	8	9	MD -0.40 (-3.09, 2.29)	Low
EXIT-25 (lower values = better score)									

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Memantine	Placebo	Summary of results	
1 (Boxer 2013)	RCT	Not serious	Not serious	N/A	Very serious[1]	8	9	MD -0.80 (-7.45, 5.85)	Low
Neuropsychiatric symptoms									
NPI (lower values = better score)									
1 (Boxer 2013)	RCT	Not serious	Not serious	N/A	Very serious[1]	8	9	MD 0.00 (-5.36, 5.36)	Low
Global assessment									
Clinician's Global Impression of Change (lower values = better score)									
1 (Boxer 2013)	RCT	Not serious	Not serious	N/A	Very serious[1]	8	9	MD 0.00 (-0.36, 0.36)	Low
Clinical Dementia Rating Sum of Boxes (lower values = better score)									
1 (Boxer 2013)	RCT	Not serious	Not serious	N/A	Very serious[1]	8	9	MD 0.90 (-0.28, 2.08)	Low
Motor function									
Unified Parkinson's disease rating scale (lower values = better score)									
1 (Boxer 2013)	RCT	Not serious	Not serious	N/A	Very serious[1]	8	9	MD 3.30 (-3.14, 9.74)	Low
Adverse events									
Serious adverse events (lower values = better score)									
1 (Boxer 2013)	RCT	Not serious	Not serious	N/A	Very serious[1]	0/8	0/9	No events in either group	Low

1. Small sample size and non-significant result.

G.7.4.4 Cognitive impairment in people with multiple sclerosis

Cholinesterase inhibitors versus placebo

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	AChEI	Placebo	Summary of results	
Cognitive outcomes – global cognition									
Selective reminding test (higher values = better score)									
2 (Krupp 2011, Maurer 2012)	RCT	Not serious	Not serious	Not serious	Serious[1]	104	97	MD 0.64 (-0.43, 1.72)	Moderate
Multiple Sclerosis Inventarium Cognition Score (lower values = better score)									

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	AChEI	Placebo	Summary of results	
1 (Maurer 2012)	RCT	Not serious	Not serious	N/A	Serious[1]	43	38	MD -0.86 (-3.17, 1.45)	Moderate
Cognitive outcomes – domain specific									
Paced Auditory Serial Addition Test 3 (higher values = better score)									
1 (Maurer 2012)	RCT	Not serious	Not serious	N/A	Serious[1]	43	38	MD 1.71 (-1.41, 4.83)	Moderate
Paced Auditory Serial Addition Test 2+3 (higher values = better score)									
1 (Krupp 2011)	RCT	Not serious	Not serious	N/A	Serious[1]	91	59	MD 0.30 (-4.08, 4.68)	Moderate
Faces Symbol Test (lower values = better score)									
1 (Maurer 2012)	RCT	Not serious	Not serious	N/A	Serious[1]	43	38	MD 0.14 (-0.36, 0.64)	Moderate
Symbol digit modalities test (higher values = better score)									
2 (Krupp 2011, Maurer 2012)	RCT	Not serious	Not serious	Not serious	Serious[1]	104	97	MD -1.40 (-3.33, 0.53)	Moderate
Depression									
Montgomery-Asberg Depression Rating Scale (lower values = better score)									
1 (Maurer 2012)	RCT	Not serious	Not serious	N/A	Serious[1]	43	38	MD -1.58 (-3.66, 0.50)	Moderate
Adverse events									
Any adverse events (lower values = better score)									
1 (Maurer 2012)	RCT	Not serious	Not serious	N/A	Serious[2]	35/45	27/41	RR 1.18 (0.90, 1.55)	Moderate
Serious adverse events (lower values = better score)									
2 (Krupp 2011, Maurer 2012)	RCT	Not serious	Not serious	Not serious	Very serious[3]	3/106	6/100	RR 0.46 (0.12, 1.70)	Low
Discontinuation due to adverse events (lower values = better score)									
1 (Maurer 2012)	RCT	Not serious	Not serious	N/A	Very serious[3]	8/45	3/41	RR 2.43 (0.69, 8.55)	Low
MS relapse									
1 (Maurer 2012)	RCT	Not serious	Not serious	N/A	Very serious[3]	4/45	6/41	RR 0.61 (0.18, 2.00)	Low

1. Non-significant result.
2. 95% CI crosses one line of a defined MID interval
3. 95% CI crosses two lines of a defined MID interval

Memantine versus placebo

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Memantine	Placebo	Summary of results	
Cognitive outcomes - domain specific									
Paced Auditory Serial Addition Test (higher values = better score)									
1 (Saint-Paul 2016)	RCT	Not serious	Not serious	N/A	Serious[1]	31	31	MD 0.70 (-6.51, 5.11)	Moderate
Multiple sclerosis progression									
Expanded Disability Status Scale (lower values = better score)									
1 (Saint-Paul 2016)	RCT	Not serious	Not serious	N/A	Serious[1]	34	34	MD -0.47 (-1.08, 0.12)	Moderate
Adverse events									
Any adverse events (lower values = better score)									
1 (Saint-Paul 2016)	RCT	Not serious	Not serious	N/A	Not serious	36/48	8/38	RR 3.56 (1.88, 6.74)	High
Discontinuation due to adverse events (lower values = better score)									
1 (Saint-Paul 2016)	RCT	Not serious	Not serious	N/A	Very serious[2]	8/50	2/43	RR 3.44 (0.77, 15.34)	Low

1. Non-significant result.
2. 95% CI crosses two lines of a defined MID interval

Network-meta analyses

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Active	Placebo	Summary of results	
Cognitive outcomes – domain specific									
Paced Auditory Serial Addition Test (higher values = better score)									
2 (Maurer 2012, Saint-Paul 2016)	RCT	Not serious	Not serious	N/A	Serious[1]	74	69	See appendix H	Moderate
Adverse events									
Any adverse events (lower values = better score)									
2 (Maurer 2012, Saint-Paul 2016)	RCT	Not serious	Not serious	N/A	Not serious	93	79	See appendix H	High
Discontinuation due to adverse events (lower values = better score)									
2 (Maurer 2012, Saint-Paul 2016)	RCT	Not serious	Not serious	N/A	Serious[1]	93	79	See appendix H	Moderate

1. Analysis could not differentiate any treatment groups.

G.7.4.5 Huntington's disease

Cholinesterase inhibitors versus placebo

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	AChIE	Placebo	Summary of results	
Cognitive outcomes- domain specific									
Symbol Digit Modalities Test score (higher values = better score)									
1 (Sesok 2014)	RCT	Not serious	Not serious	N/A	Very serious[1]	11	6	MD 15.17 (-28.82, 59.16)	Low
Tower of London total moves score (higher values = better score)									
1 (Sesok 2014)	RCT	Not serious	Not serious	N/A	Very serious[1]	11	6	MD 20.18 (-10.53, 50.89)	Low
Tower of London total time score (lower values = better score)									
1 Sesok 2014)	RCT	Not serious	Not serious	N/A	Serious[2]	11	6	MD 268.47 (118.84, 418.10)	Moderate
Rey Complex Figure Test – delayed recall (higher values = better score)									
1 (Sesok 2014)	RCT	Not serious	Not serious	N/A	Very serious[1]	11	6	MD -2.86 (-10.90, 5.18	Low
Rey Complex Figure Test - immediate recall (higher values = better score)									
1 (Sesok 2014)	RCT	Not serious	Not serious	N/A	Very serious[1]	11	6	MD -3.77 (-11.92, 4.38)	Low
Ruff Figural Fluency Test - unique designs score (higher values = better score)									
1 (Sesok 2014)	RCT	Not serious	Not serious	N/A	Very serious[1]	11	6	MD -3.03 (-31.17, 25.11)	Low

1. Small sample size and non-significant result.
2. Small sample size.

G.8 Drugs that may worsen cognitive decline

G.8.1 Drugs that may cause cognitive decline

- What drugs that may worsen cognitive decline are commonly prescribed in people diagnosed with dementia?
- What are the most effective tools to identify whether drugs may be the cause of cognitive decline in someone suspected of having dementia?

No GRADE or CERQual tables were produced for this review question

G.9 Non-pharmacological interventions for dementia

G.9.1 Non-pharmacological interventions for people living with dementia

- What are the most effective non-pharmacological interventions for supporting cognitive functioning in people living with dementia?
- What are the most effective non-pharmacological interventions for supporting functional ability in people living with dementia?
- What are the most effective non-pharmacological interventions to support wellbeing in people living with dementia?
- What are the most effective methods of supporting people living with dementia to reduce harm and stay independent?

G.9.1.1 Cognitive stimulation therapy

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
Cognition: MMSE (post-intervention) – higher numbers favour intervention							
20	Not serious	Serious[1]	Not serious	Not serious	1,341	MD 1.76 (1.01, 2.51)	Moderate
Cognition: MMSE (follow-up) – higher numbers favour intervention							
2	Not serious	Serious[1]	Not serious	Serious[2]	77	MD 2.99 (-2.33, 8.31)	Low
Cognition: all measures (post-intervention) – higher numbers favour intervention							
25	Not serious	Serious[1]	Not serious	Not serious	1,398	SMD 0.44 (0.27, 0.62)	Moderate
Cognition: all measures (follow-up) – higher numbers favour intervention							
4	Not serious	Not serious	Not serious	Serious[3]	106	SMD 0.42 (0.03, 0.81)	Moderate
ADL: ADCS-ADL (post-intervention) – higher numbers favour intervention							
1 (Orrell 2014)	Not serious	N/A	Not serious	Serious[2]	236	MD 0.94 (-2.04, 3.92)	Moderate
ADL: all measures (post-intervention) – higher numbers favour intervention							
8	Not serious	Not serious	Not serious	Serious[3]	784	SMD 0.13 (-0.01, 0.27)	Moderate
Clinical dementia rating scale (post-intervention) – lower numbers favour intervention							
2	Serious[4]	Not serious	Not serious	Serious[2]	73	MD -0.23 (-0.53, 0.07)	Low
Behavioural and psychological symptoms: NPI (post-intervention) – lower numbers favour intervention							
3	Not serious	Serious[1]	Not serious	Serious[2]	644	MD -0.12 (-2.10, 1.85)	Low
Behavioural and psychological symptoms: NPI (follow-up) – lower numbers favour intervention							

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
1 (Chapman 2004)	Not serious	N/A	Not serious	Serious[2]	54	MD -4.44 (-12.35, 3.47)	Moderate
Behavioural and psychological symptoms: all measures (post-intervention) – higher numbers favour intervention							
8	Not serious	Serious[1]	Not serious	Serious[3]	921	SMD 0.05 (-0.16, 0.26)	Low
Behavioural and psychological symptoms: all measures (follow-up) – higher numbers favour intervention							
2	Not serious	Not serious	Not serious	Serious[3]	64	SMD 0.37 (-0.13, 0.87)	Moderate
Depression: Cornell scale for depression in dementia (post-intervention) – lower numbers favour intervention							
3	Not serious	Serious[1]	Not serious	Serious[2]	194	MD -0.30 (-2.11, 1.51)	Low
Depression: all measures (post-intervention) – lower numbers favour intervention							
12	Not serious	Not serious	Not serious	Serious[3]	746	SMD 0.05 (-0.10, 0.19)	Moderate
Quality of life: QoL-AD (post-intervention) – higher numbers favour intervention							
10	Not serious	Serious[1]	Not serious	Serious[2]	885	MD 0.47 (-0.17, 1.10)	Low
Quality of life: QoL-AD (follow-up) – higher numbers favour intervention							
2	Not serious	Not serious	Not serious	Not serious	290	MD 1.87 (0.29, 3.44)	High
Quality of life: EQ-5D (post-intervention) – higher numbers favour intervention							
1 (Yamanaka 2013)	Not serious	N/A	Not serious	Very serious[5]	50	MD 0.01 (-0.12, 0.14)	Low
Quality of life: all measures (post-intervention) – higher numbers favour intervention							
11	Not serious	Serious[1]	Not serious	Serious[3]	895	SMD 0.10 (-0.03, 0.23)	Low
Quality of life: all measures (follow-up) – higher numbers favour intervention							
3	Not serious	Not serious	Not serious	Serious[3]	300	SMD 0.26 (0.03, 0.49)	Moderate
Carer burden: all measures (post-intervention) – higher numbers favour intervention							
4	Not serious	Not serious	Not serious	Not serious	435	SMD 0.00 (-0.18, 0.19)	High

1. $i^2 > 40\%$
2. Non-significant result
3. 95% CI crosses 1 line of a defined MID interval
4. No details of randomisation method or assessor blinding reported
5. Non-significant result and small sample size

G.9.1.2 Cognitive training

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
Cognition: MMSE (post-intervention) – higher numbers favour intervention							
9	Not serious	Serious[1]	Not serious	Serious[2]	252	MD 1.31 (-1.36, 3.98)	Low
Cognition: MMSE (follow-up) – higher numbers favour intervention							
2	Serious[3]	Serious[1]	Not serious	Very serious[4]	24	MD 0.96 (-3.19, 5.11)	Very low
Cognition: all measures (post-intervention) – higher numbers favour intervention							
12	Not serious	Serious[1]	Not serious	Serious[5]	608	SMD 0.36 (-0.00, 0.73)	Low
Cognition: all measures (follow-up) – higher numbers favour intervention							
6	Not serious	Not serious	Not serious	Serious[5]	385	SMD 0.04 (-0.16, 0.24)	Moderate
ADL: all measures (post-intervention) – higher numbers favour intervention							
6	Not serious	Not serious	Not serious	Serious[5]	444	SMD 0.12 (-0.07, 0.31)	Moderate
ADL: all measures (follow-up) – higher numbers favour intervention							
5	Not serious	Not serious	Not serious	Very serious[6]	366	SMD -0.00 (-0.21, 0.20)	Low
Behavioural and psychological symptoms: NPI (post-intervention) – lower numbers favour intervention							
1 (Amieva 2016)	Not serious	N/A	Not serious	Serious[2]	292	MD 1.81 (-1.57, 5.19)	Moderate
Behavioural and psychological symptoms: NPI (follow-up) – lower numbers favour intervention							
1 (Amieva 2016)	Not serious	N/A	Not serious	Serious[2]	233	MD 3.73 (-0.38, 7.84)	Moderate
Behavioural and psychological symptoms: all measures (post-intervention) – higher numbers favour intervention							
1 (Amieva 2016)	Not serious	N/A	Not serious	Serious[5]	292	SMD -0.12 (-0.35, 0.11)	Moderate
Behavioural and psychological symptoms: all measures (follow-up) – higher numbers favour intervention							
1 (Amieva 2016)	Not serious	N/A	Not serious	Serious[5]	233	SMD -0.23 (-0.49, 0.03)	Moderate
Depression: Cornell scale for depression in dementia (post-intervention) – higher numbers favour intervention							
1 (Bergamaschi 2013)	Serious[3]	N/A	Not serious	Very serious[4]	32	MD -1.51 (-5.99, 2.77)	Very low
Depression: all measures (post-intervention) – higher numbers favour intervention							
7	Not serious	Serious[1]	Not serious	Serious[5]	392	SMD -0.03 (-0.23, 0.17)	Low
Depression: all measures (follow-up) – higher numbers favour intervention							

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
1 (Galante 2007)	Very serious[7]	N/A	Not serious	Very serious[6]	11	SMD 0.05 (-1.18, 1.28)	Very low
Quality of life: QoL-AD (post-intervention) – higher numbers favour intervention							
1 (Amieva 2016)	Not serious	N/A	Not serious	Serious[2]	292	MD -0.87 (-1.93, 0.19)	Moderate
Quality of life: QoL-AD (post-intervention) – higher numbers favour intervention							
1 (Amieva 2016)	Not serious	N/A	Not serious	Serious[2]	233	MD -0.93 (-2.10, 0.24)	Moderate
Quality of life: all measures (post-intervention) – higher numbers favour intervention							
1 (Amieva 2016)	Not serious	N/A	Not serious	Serious[5]	292	SMD -0.19 (-0.42, 0.04)	Moderate
Quality of life: all measures (follow-up) – higher numbers favour intervention							
1 (Amieva 2016)	Not serious	N/A	Not serious	Serious[5]	233	SMD -0.20 (-0.46, 0.05)	Moderate
Carer burden: all measures (post-intervention) – higher numbers favour intervention							
3	Not serious	Not serious	Not serious	Serious[5]	372	SMD -0.09 (-0.29, 0.12)	Moderate
Carer burden: all measures (follow-up) – higher numbers favour intervention							
1 (Amieva 2016)	Not serious	N/A	Not serious	Serious[5]	233	SMD -0.22 (-0.48, 0.04)	Moderate

1. $i^2 > 40\%$
2. Non-significant result
3. No details of randomisation method or assessor blinding reported
4. Non-significant result and small sample size
5. 95% CI crosses 1 line of a defined MID interval
6. 95% CI crosses 2 lines of a defined MID interval
7. No details of randomisation method or assessor blinding reported. Post-hoc exclusion of participants for 'poor compliance'

G.9.1.3 Cognitive rehabilitation

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
Cognition: MMSE (post-intervention) – higher numbers favour intervention							
1 (Seyun 2015)	Serious[1]	N/A	Not serious	Not serious	43	MD 1.00 (0.32, 1.68)	Moderate
Cognition: all measures (post-intervention) – higher numbers favour intervention							
2	Not serious	Serious[2]	Not serious	Very serious[3]	328	SMD 0.42 (-0.36, 1.19)	Very low

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
Cognition: all measures (follow-up) – higher numbers favour intervention							
1 (Amieva 2016)	Not serious	N/A	Not serious	Very serious[3]	230	SMD -0.04 (-0.30, 0.22)	Low
ADL: all measures (post-intervention) – higher numbers favour intervention							
4	Not serious	Serious[2]	Not serious	Serious[4]	812	SMD 0.44 (-0.09, 0.96)	Low
ADL: all measures (follow-up) – higher numbers favour intervention							
2	Not serious	Serious[2]	Not serious	Very serious[3]	646	SMD 0.62 (-0.05, 1.30)	Very low
Behavioural and psychological symptoms: NPI (post-intervention) – lower numbers favour intervention							
2	Not serious	Not serious	Not serious	Serious[5]	302	MD 2.20 (-1.39, 5.79)	Moderate
Behavioural and psychological symptoms: NPI (follow-up) – lower numbers favour intervention							
2	Not serious	Serious[2]	Not serious	Serious[5]	247	MD 0.09 (-8.74, 10.54)	Low
Behavioural and psychological symptoms: all measures (post-intervention) – higher numbers favour intervention							
2	Not serious	Not serious	Not serious	Serious[4]	302	SMD -0.14 (-0.36, 0.09)	Moderate
Behavioural and psychological symptoms: all measures (follow-up) – higher numbers favour intervention							
2	Not serious	Serious[2]	Not serious	Very serious[3]	247	SMD -0.07 (-0.81, 0.68)	Very low
Depression: all measures (post-intervention) – higher numbers favour intervention							
3	Not serious	Serious[2]	Not serious	Serious[4]	770	SMD -0.11 (-0.35, 0.13)	Low
Depression: all measures (follow-up) – higher numbers favour intervention							
3	Not serious	Not serious	Not serious	Not serious	670	SMD -004 (-0.19, 0.11)	High
Quality of life: QoL-AD (post-intervention) – higher numbers favour intervention							
3	Not serious	Serious[2]	Not serious	Serious[5]	369	MD 0.80 (-1.59, 3.19)	Moderate
Quality of life: QoL-AD (follow-up) – higher numbers favour intervention							
2	Not serious	Not serious	Not serious	Serious[5]	258	MD -0.15 (-1.29, 1.00)	Moderate
Quality of life: all measures (post-intervention) – higher numbers favour intervention							
5	Not serious	Not serious	Not serious	Not serious	831	SMD 0.02 (-0.11, 0.16)	High
Quality of life: all measures (follow-up) – higher numbers favour intervention							
4	Not serious	Not serious	Not serious	Not serious	692	SMD 0.01 (-0.14, 0.16)	High

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
Carer burden: all measures (post-intervention) – higher numbers favour intervention							
4	Not serious	Not serious	Not serious	Not serious	754	SMD 0.04 (-0.10, 0.18)	High
Carer burden: all measures (follow-up) – higher numbers favour intervention							
4	Not serious	Not serious	Not serious	Not serious	674	SMD -0.01 (-0.16, 0.14)	High

1. No details of randomisation method or assessor blinding reported
2. $i^2 > 40\%$
3. 95% CI crosses 2 lines of a defined MID interval
4. 95% CI crosses 1 line of a defined MID interval
5. Non-significant result

G.9.1.4 Self-management groups

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
Cognition: all measures (post-intervention) – higher numbers favour intervention							
1 (Laakonen 2016)	Not serious	N/A	Not serious	Serious[2]	134	SMD -0.28 (-0.62, 0.06)	Moderate
Depression: all measures (post-intervention) – lower numbers favour intervention							
1 (Logsdon 2010)	Serious[4]	N/A	Not serious	Serious[2]	134	SMD -0.26 (-0.62, 0.10)	Low
Depression: all measures (follow-up) – lower numbers favour intervention							
1 (Quinn 2016)	Not serious	N/A	Not serious	Very Serious[3]	23	SMD 0.30 (-0.52, 1.12)	Low
Quality of life: QoL-AD (post-intervention) – higher numbers favour intervention							
1 (Logsdon 2010)	Serious[4]	N/A	Not serious	Serious[1]	134	MD 1.67 (-0.44, 3.78)	Low
Quality of life: EQ-5D (post-intervention) – higher numbers favour intervention							
1 (Quinn 2016)	Not serious	N/A	Not serious	Serious[1]	23	MD 0.05 (-0.04, 0.14)	Moderate
Quality of life: EQ-5D (follow-up) – higher numbers favour intervention							
1 (Quinn 2016)	Not serious	N/A	Not serious	Serious[1]	23	MD -0.04 (-0.15, 0.07)	Moderate
Quality of life: all measures (post-intervention) – higher numbers favour intervention							
3	Not serious	Not serious	Not serious	Serious[2]	291	SMD 0.24 (-0.00, 0.47)	Moderate
Quality of life: all measures (follow-up) – higher numbers favour intervention							

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
1 (Quinn 2016)	Not serious	N/A	Not serious	Very Serious[3]	23	SMD -0.29 (-1.11, 0.54)	Low

1. Non-significant result
2. 95% CI crosses 1 line of a defined MID interval
3. 95% CI crosses 2 lines of a defined MID interval
4. Outcomes assessors not blinded

G.9.1.5 Reminiscence therapy

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
Cognition: MMSE (post-intervention) – higher numbers favour intervention							
8	Not serious	Serious[1]	Not serious	Not serious	491	MD 1.55 (0.77, 2.33)	Moderate
Cognition: MMSE (follow-up) – higher numbers favour intervention							
1 (Tadaka 2007)	Serious[2]	N/A	Not serious	Not serious	50	MD 1.49 (0.57, 2.40)	Moderate
Cognition: all measures (post-intervention) – higher numbers favour intervention							
9	Not serious	Serious[1]	Not serious	Serious[4]	782	SMD 0.28 (0.14, 0.42)	Low
Cognition: all measures (follow-up) – higher numbers favour intervention							
2	Serious[2]	Serious[1]	Not serious	Very serious[5]	277	SMD 0.35 (-0.64, 1.33)	Very low
ADCS-ADL: all measures (post-intervention) – higher numbers favour intervention							
1 (Deponte 2007)	Not serious	Not serious	Not serious	Serious[3]	18	MD -2.40 (-6.93, 2.13)	Moderate
ADL: all measures (post-intervention) – higher numbers favour intervention							
4	Not serious	Not serious	Not serious	Not serious	993	SMD -0.00 (-0.13, 0.12)	High
ADL: all measures (follow-up) – higher numbers favour intervention							
2	Not serious	Serious[1]	Not serious	Very serious[5]	577	SMD -0.01 (-0.35, 0.34)	Very low
BPSD: NPI (post-intervention) – lower numbers favour intervention							
3	Not serious	Not serious	Not serious	Serious[3]	614	MD 0.28 (-2.05, 2.61)	Moderate
BPSD: NPI (follow-up) – lower numbers favour intervention							
1 (Amieva 2016)	Not serious	N/A	Not serious	Serious[3]	227	MD 1.71 (-2.42, 5.84)	Moderate
BPSD: all measures (post-intervention) – lower numbers favour intervention							

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
5	Not serious	Not serious	Not serious	Not serious	714	SMD 0.04 (-0.11, 0.19)	High
BPSD: all measures (follow-up) – lower numbers favour intervention							
1 (Amieva 2016)	Not serious	N/A	Not serious	Serious[4]	227	SMD 0.11 (-0.15, 0.37)	Moderate
Depression: CSDD (post-intervention) – lower numbers favour intervention							
3	Not serious	Serious[1]	Not serious	Serious[3]	537	MD -1.51 (-3.70, 0.67)	Low
Depression: CSDD (follow-up) – lower numbers favour intervention							
1 (Woods 2016)	Not serious	N/A	Not serious	Serious[3]	350	MD 0.38 (-0.85, 1.61)	Moderate
Depression: all measures (post-intervention) – lower numbers favour intervention							
8	Not serious	Serious[1]	Not serious	Very serious[5]	1,432	SMD -0.15 (-0.38, 0.07)	Very low
Depression: all measures (follow-up) – lower numbers favour intervention							
2	Not serious	Not serious	Not serious	Serious[4]	577	SMD 0.04 (-0.12, 0.21)	Moderate
Quality of life: QoL-AD (post-intervention) – higher numbers favour intervention							
4	Not serious	Serious[1]	Not serious	Serious[3]	998	MD 0.53 (-0.97, 2.02)	Low
Quality of life: QoL-AD (follow-up) – higher numbers favour intervention							
2	Not serious	Not serious	Not serious	Serious[3]	577	MD 0.19 (-0.73, 1.11)	Moderate
Quality of life: EQ-5D (post-intervention) – higher numbers favour intervention							
2	Not serious	Not serious	Not serious	Serious[3]	684	MD 0.01 (-0.03, 0.05)	Moderate
Quality of life: EQ-5D (follow-up) – higher numbers favour intervention							
1 (Woods 2016)	Not serious	N/A	Not serious	Serious[3]	350	MD 0.00 (-0.05, 0.06)	Moderate
Quality of life: all measures (post-intervention) – higher numbers favour intervention							
5	Not serious	Serious[1]	Not serious	Serious[4]	1,071	SMD 0.09 (-0.12, 0.30)	Low
Quality of life: all measures (follow-up) – higher numbers favour intervention							
3	Not serious	Not serious	Not serious	Serious[4]	650	SMD 0.03 (-0.13, 0.18)	Moderate
Agitation: CMAI (post-intervention) – lower numbers favour intervention							
1 (Eritz 2015)	Not serious	N/A	Not serious	Serious[3]	73	MD -1.07 (-7.52, 5.38)	Moderate
Agitation: CMAI (follow-up) – lower numbers favour intervention							

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
1 (Eritz 2015)	Not serious	N/A	Not serious	Serious[3]	73	MD 0.96 (-12.10, 14.302	Moderate
Agitation: all measures (post-intervention) – lower numbers favour intervention							
1 (Eritz 2015)	Not serious	N/A	Not serious	Very serious[5]	73	SMD -0.17 (-0.53, 0.39)	Low
Agitation: all measures (follow-up) – lower numbers favour intervention							
1 (Eritz 2015)	Not serious	N/A	Not serious	Very serious[5]	73	SMD 0.03 (-0.43, 0.49)	Low
Carer burden: all measures (post-intervention) – lower numbers favour intervention							
2	Not serious	Not serious	Not serious	Serious[4]	580	SMD -0.03 (-0.20, 0.14)	Moderate
Carer burden: all measures (follow-up) – lower numbers favour intervention							
1 (Amieva 2016)	Not serious	N/A	Not serious	Very serious[5]	227	SMD 0.00 (-0.26, 0.26)	Low

1. $i^2 > 40\%$
2. No details of randomisation method or assessor blinding reported
3. Non-significant result
4. 95% CI crosses 1 line of a defined MID interval
5. 95% CI crosses 2 lines of a defined MID interval

G.9.1.6 Occupational therapy

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
ADL: all measures (post-intervention) – higher numbers favour intervention							
2	Not serious	Not serious	Not serious	Very serious[2]	313	SMD 0.14 (-0.24, 0.53)	Low
ADL: all measures (follow-up) – higher numbers favour intervention							
1 (Voigt Radlof 2011)	Not serious	N/A	Not serious	Serious[1]	104	SMD -0.19 (-0.58, 0.19)	Moderate
Depression: CSDD (post-intervention) – lower numbers favour intervention							
3	Not serious	Not serious	Not serious	Not serious	266	MD -2.29 (-3.47, -1.10)	High
Depression: CSDD (follow-up) – lower numbers favour intervention							
2	Not serious	Serious[3]	Not serious	Not serious	210	MD -2.79 (-4.41, -1.18)	Low
Depression: all measures (post-intervention) – lower numbers favour intervention							

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
3	Not serious	Not serious	Not serious	Serious[1]	266	SMD -0.44 (-0.69, -0.20)	Moderate
Depression: all measures (follow-up) – lower numbers favour intervention							
2	Not serious	Not serious	Not serious	Serious[1]	210	SMD -0.45 (-0.76, -0.18)	Moderate
Quality of life: QoL-AD (post-intervention) – higher numbers favour intervention							
2	Not serious	Not serious	Not serious	Serious[4]	265	MD 0.10 (0.01, 0.19)	Moderate
Quality of life: all measures (post-intervention) – higher numbers favour intervention							
4	Not serious	Serious[3]	Not serious	Serious[1]	491	SMD 0.50 (0.09, 0.91)	Low
Quality of life: all measures (follow-up) – higher numbers favour intervention							
2	Not serious	Serious[3]	Not serious	Serious[1]	226	SMD 0.68 (-0.12, 1.48)	Low
Agitation: all measures (post-intervention) – lower numbers favour intervention							
1 (Gitlin 2010)	Not serious	N/A	Not serious	Very serious[2]	209	SMD 0.00 (-0.27, 0.27)	Low
Carer burden: ZBI (post-intervention) – lower numbers favour intervention							
1 (Gitlin 2008)	Serious[5]	N/A	Not serious	Serious[4]	56	SMD 0.00 (-4.91, 4.91)	Low
Carer burden: all measures (post-intervention) – lower numbers favour intervention							
2	Serious[5]	Serious[3]	Not serious	Serious[1]	265	SMD 0.27 (-0.13, 0.67)	Very low

1. 95% CI crosses 1 line of a defined MID interval
2. 95% CI crosses 2 lines of a defined MID interval
3. $i^2 > 40\%$
4. Non-significant result
5. No details of randomisation method or assessor blinding reported

G.9.1.7 Psychotherapy

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
Cognition: MMSE (post-intervention) – higher numbers favour intervention							
2	Not serious	Not serious	Not serious	Serious[1]	95	MD -1.41 (-2.91, 0.10)	Moderate
Cognition: MMSE (follow-up) – higher numbers favour intervention							
2	Not serious	Not serious	Not serious	Serious[1]	92	MD -0.82 (-2.47, 0.84)	Moderate

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
Cognition: all measures (post-intervention) – higher numbers favour intervention							
2	Not serious	Not serious	Not serious	Serious[2]	95	SMD -0.36 (-0.77, 0.04)	Moderate
Cognition: all measures (follow-up) – higher numbers favour intervention							
2	Not serious	Not serious	Not serious	Very serious[3]	92	SMD -0.18 (-0.59, 0.23)	Low
ADL: all measures (post-intervention) – higher numbers favour intervention							
1 (Burns 2005)	Not serious	N/A	Not serious	Very serious[3]	40	SMD -0.37 (-1.00, 0.26)	Low
ADL: all measures (follow-up) – higher numbers favour intervention							
1 (Burns 2005)	Not serious	N/A	Not serious	Very serious[3]	40	SMD -0.17 (-0.79, 0.45)	Low
Depression: CSDD (post-intervention) – lower numbers favour intervention							
2	Not serious	Not serious	Not serious	Serious[1]	95	MD -0.86 (-2.27, 0.54)	Moderate
Depression: CSDD (follow-up) – lower numbers favour intervention							
2	Not serious	Not serious	Not serious	Serious[1]	92	MD -1.16 (-2.54, 0.22)	Moderate
Depression: all measures (post-intervention) – lower numbers favour intervention							
3	Not serious	Not serious	Not serious	Serious[2]	125	SMD -0.39 (-0.75, -0.04)	Moderate
Depression: all measures (follow-up) – lower numbers favour intervention							
2	Not serious	Not serious	Not serious	Serious[2]	92	SMD -0.32 (-0.73, 0.10)	Moderate
Quality of life: QoL-AD (post-intervention) – higher numbers favour intervention							
1 (Marshall 2014)	Not serious	N/A	Not serious	Serious[1]	55	MD 2.20 (-1.42, 5.82)	Moderate
Quality of life: QoL-AD (follow-up) – higher numbers favour intervention							
1 (Marshall 2014)	Not serious	N/A	Not serious	Serious[1]	52	MD 0.30 (-2.99, 3.59)	Moderate
Quality of life: all measures (post-intervention) – higher numbers favour intervention							
1 (Marshall 2014)	Not serious	N/A	Not serious	Very serious[3]	55	SMD 0.32 (-0.22, 0.85)	Low
Quality of life: all measures (follow-up) – higher numbers favour intervention							
1 (Marshall 2014)	Not serious	N/A	Not serious	Very serious[3]	52	SMD 0.05 (-0.50, 0.59)	Low

1. Non-significant result
2. 95% CI crosses 1 line of a defined MID interval

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
3. 95% CI crosses 2 lines of a defined MID interval							

G.9.1.8 Exercise

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

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

1. $i^2 > 40\%$
2. Evidence of selective outcome reporting

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

G.9.1.9 Nutrition

Ginkgo biloba versus placebo (Alzheimer's disease)

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
Cognition: MMSE (post-intervention) – higher numbers favour intervention							
1 (Mazza 2006)	Not serious	N/A	Not serious	Serious[1]	51	MD 0.85 (-2.39, 4.09)	Moderate
Cognition: all measures (post-intervention) – higher numbers favour intervention							
4	Not serious	Serious[2]	Not serious	Serious[3]	619	SMD 0.08 (-0.19, 0.35)	Low
ADL: all measures (post-intervention) – higher numbers favour intervention							
1 (Schneider 2005)	Not serious	N/A	Not serious	Serious[1]	343	MD 0.00 (-0.21, 0.21)	Moderate
Global assessment: MMSE (post-intervention) – higher numbers favour intervention							
1 (Le Bars 1997)	Not serious	N/A	Not serious	Serious[1]	236	MD 0.00 (-0.26, 0.26)	Moderate
1. Non-significant result 2. $i^2 > 40\%$ 3. 95% CI crosses 1 line of a defined MID interval							

Ginkgo biloba versus placebo (Alzheimer's disease or vascular dementia)

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
Cognition: MMSE (post-intervention) – higher numbers favour intervention							
6	Not serious	Serious[1]	Not serious	Serious[2]	1,922	SMD 0.60 (0.06, 1.13)	Low

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
ADL: all measures (post-intervention) – higher numbers favour intervention							
6	Not serious	Serious[1]	Not serious	Serious[2]	1,922	SMD 0.41 (0.11, 0.71)	Low
BPSD: NPI (post-intervention) – lower numbers favour intervention							
4	Not serious	Serious[1]	Not serious	Not serious	1,598	MD -3.88 (-7.63, -0.14)	Moderate
BPSD: all measures (post-intervention) – lower numbers favour intervention							
4	Not serious	Serious[1]	Not serious	Serious[2]	1,598	SMD -0.67 (-1.31, -0.03)	Low
Global assessment: all measures (post-intervention) – lower numbers favour intervention							
4	Not serious	Serious[1]	Not serious	Serious[2]	1,597	SMD 0.74 (0.14, 1.33)	Low
Quality of life: all measures (post-intervention) – lower numbers favour intervention							
2	Not serious	Not serious	Not serious	Serious[2]	806	SMD 0.24 (0.11, 0.38)	Moderate

1. $i^2 > 40\%$
2. 95% CI crosses 1 line of a defined MID interval

Omega-3 fatty acids (DHA and EPA) versus placebo

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
Cognition: MMSE (post-intervention) - higher numbers favour intervention							
3	Not serious	Not serious	Not serious	Serious[1]	604	MD 0.17 (-0.38, 0.72)	Moderate
ADL: ADCS-ADL (post-intervention) - higher numbers favour intervention							
1 (Quinn 2010)	Not serious	N/A	Not serious	Serious[1]	400	MD 1.08 (-1.70, 3.86)	Moderate
ADL: all measures (post-intervention) - higher numbers favour intervention							
2	Not serious	Not serious	Not serious	Serious[2]	426	SMD 0.04 (-0.15, 0.24)	Moderate
BPSD: NPI (post-intervention) - lower numbers favour intervention							
1 (Quinn 2010)	Not serious	N/A	Not serious	Serious[1]	400	MD -2.16 (-5.42, 1.10)	Moderate
Dementia severity: CDR (post-intervention) - lower numbers favour intervention							
2	Not serious	Not serious	Not serious	Serious[1]	578	MD -0.07 (-0.63, 0.48)	Moderate

1. Non-significant result
2. 95% CI crosses 1 line of a defined MID interval

Souvenaid versus placebo

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
Cognition: MMSE (post-intervention) - higher numbers favour intervention							
1 (Scheltens 2010)	Not serious	N/A	Not serious	Serious[1]	195	MD 0.30 (-0.56, 1.16)	Moderate
Cognition: all measures (post-intervention) - higher numbers favour intervention							
3	Not serious	Serious[2]	Not serious	Serious[3]	879	SMD 0.10 (-0.12, 0.32)	Low
ADL: ADCS-ADL (post-intervention) - higher numbers favour intervention							
3	Not serious	Not serious	Not serious	Serious[1]	651	MD 0.13 (-1.32, 1.58)	Moderate
Quality of life: QoL-AD (post-intervention) - higher numbers favour intervention							
1 (Scheltens 2010)	Not serious	N/A	Not serious	Serious[1]	200	MD -0.40 (-1.59, 0.79)	Moderate
Dementia severity: CDR (post-intervention) - lower numbers favour intervention							
1 (Shah 2013)	Not serious	N/A	Not serious	Serious[1]	450	MD 0.08 (-0.28, 0.44)	Moderate

1. Non-significant result
2. $i^2 > 40\%$
3. 95% CI crosses 1 line of a defined MID interval

Huperzine A versus placebo or no treatment

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
Cognition: MMSE (post-intervention) - higher numbers favour Huperzine							
7	Very serious[1]	Serious[3]	Not serious	Not serious	648	MD 2.80 (1.61, 3.99)	Very low
ADL: ADCS-ADL (post-intervention) - higher numbers favour Huperzine							
1 (Rafii 2011)	Not serious	N/A	Not serious	Serious[2]	210	MD 1.63 (-0.84, 4.09)	Moderate
ADL: all measures (post-intervention) - higher numbers favour Huperzine							
7	Very serious[1]	Serious[3]	Not serious	Not serious	648	SMD 0.54 (0.23, 0.85)	Very low
Dementia severity: CDR (post-intervention) - higher numbers favour Huperzine							
1 (Yang 2003)	Very serious[1]	N/A	Not serious	Not serious	65	MD -0.80 (-0.95, -0.65)	Low
BPSD:NPI (post-intervention) – higher numbers favour Huperzine							
1 (Rafii 2011)	Not serious	N/A	Not serious	Serious[2]	210	MD 0.15 (-2.35, 2.66)	Moderate

1. Individual studies at high risk of bias, and data not available from some studies only reported in Chinese
2. Non-significant result
3. $i^2 > 40\%$

Tailored nutritional guidance versus normal community care

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
Quality of life: 15D (post-intervention) – higher numbers favour tailored nutritional guidance							
1 (Suominen 2015)	Serious[1]	N/A	Not serious	Not serious	78	MD 0.04 (0.01, 0.07)	Moderate

1. Intention to treat analysis not carried out

Multivitamins versus placebo

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
Cognition: MMSE (post-intervention) – higher numbers favour tailored nutritional guidance							
1 (Sun 2007)	Serious[1]	N/A	Not serious	Serious[2]	89	MD -0.26 (-2.16, 1.64)	Low
ADL: Barthel Index (post-intervention) – higher numbers favour tailored nutritional guidance							
1 (Sun 2007)	Serious[1]	N/A	Not serious	Serious[2]	89	MD -0.14 (-0.91, 0.63)	Low

1. No details of randomisation method or assessor blinding reported
2. Non-significant result

Vitamin E versus placebo

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
Cognition: MMSE (post-intervention) – higher numbers favour vitamin E							
1 (Dysken 2014)	Not serious	Serious[2]	Not serious	Serious[1]	561	MD 0.22 (-0.13, 0.87)	Moderate
ADL:ADCS-ADL (post-intervention) – higher numbers favour vitamin E							
1 (Dysken 2014)	Not serious	Not serious	Not serious	Serious[1]	561	MD 1.46 (-1.84, 4.76)	Moderate
BPSD:NPI (post-intervention) – higher numbers favour vitamin E							
1 (Dysken 2014)	Not serious	Not serious	Not serious	Serious[1]	561	MD -0.77 (-2.74, 1.19)	Moderate

1. Not serious

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
2. i^2 > 40%							

Folic Acid, B12 and B6 versus placebo

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
Cognition: MMSE (post-intervention) – higher numbers favour intervention							
1 (Aisen 2008)	Not serious	N/A	Not serious	Serious[1]	409	MD -0.43 (-1.32, 0.46)	Moderate
ADL: ADCSL-ADL (post-intervention) – higher numbers favour intervention							
1 (Aisen 2008)	Not serious	N/A	Not serious	Serious[1]	409	MD -0.96 (-3.25, 1.33)	Moderate
Dementia severity: CDR (post-intervention) – lower numbers favour intervention							
1 (Aisen 2008)	Not serious	N/A	Not serious	Serious[1]	409	MD 0.07 (-0.41, 0.55)	Moderate
1. Non-significant result							

Folic acid, B12, Hcy, SAM, SAH and donepezil versus donepezil

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
Cognition: MMSE (post-intervention) – higher numbers favour intervention							
2	Serious[1]	Not serious	Not serious	Serious[2]	162	MD 0.26 (-1.22, 1.74)	Low
ADL: all measures (post-intervention) – higher numbers favour intervention							
2	Serious[1]	N/A	Not serious	Very serious[3]	162	SMD 0.28 (-0.38, 0.95)	Very low
1. Intention to treat analysis not carried out 2. Non-significant result 3. 95% CI crosses 2 lines of a defined MID interval							

Oral nutritional supplements versus standard dietetic advice

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
Cognition: MMSE (post-intervention) – higher numbers favour intervention							
2	Not serious	Not serious	Not serious	Serious[1]	58	MD 0.68 (-0.96, 2.31)	Moderate
Cognition: MMSE (follow-up) – higher numbers favour intervention							

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
2	Not serious	Not serious	Not serious	Serious[1]	55	MD 0.39 (-1.55, 2.33)	Moderate
ADL: all measures (post-intervention) – higher numbers favour intervention							
2	Not serious	Not serious	Not serious	Very serious[2]	115	SMD 0.07 (-0.30, 0.44)	Low
ADL: all measures (follow-up) – higher numbers favour intervention							
1 (Lauque 2004)	Not serious	N/A	Not serious	Very serious[2]	80	SMD 0.08 (-0.35, 0.51)	Low

1. Non-significant result
2. 95% CI crosses 2 lines of a defined MID interval

Whole formula diet (based on lyophilised (dried) foods) versus standard dietetic advice

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
Cognition: all measures (post-intervention) – higher numbers favour intervention							
1 (Salas-Salvado 2004)	Serious[1]	N/A	Not serious	Very serious[2]	38	SMD -0.38 (-1.04, 0.28)	Very low

1. Intention to treat analysis not carried out
2. 95% CI crosses 2 lines of a defined MID interval

Ginseng versus placebo

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
Cognition: MMSE (post-intervention) – higher numbers favour intervention							
3	Serious[1]	N/A	Not serious	Serious[2]	226	MD 0.31 (-0.52, 1.15)	Low

1. Open-label study
2. Non-significant result

Chinese herbal formula (Yishen Huazhuo decoction) and donepezil versus placebo and donepezil

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
Cognition: MMSE (post-intervention) – higher numbers favour intervention							
1 (Zhang 2015)	Not serious	N/A	Serious[1]	Serious[2]	144	MD 0.45 (-0.34, 1.24)	Low

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
Cognition: MMSE (follow-up) – higher numbers favour intervention							
1 (Zhang 2015)	Not serious	N/A	Serious[1]	Not serious	144	MD 0.97 (0.25, 1.69)	Moderate
ADL: all measures (post-intervention) – higher numbers favour intervention							
1 (Zhang 2015)	Not serious	N/A	Serious[1]	Serious[3]	144	SMD -0.01 (-0.34, 0.31)	Low
ADL: all measures (follow-up) – higher numbers favour intervention							
1 (Zhang 2015)	Not serious	N/A	Serious[1]	Serious[2]	144	SMD -0.23 (-0.56, 0.10)	Low
BPSD: NPI (post-intervention) – lower numbers favour intervention							
1 (Zhang 2015)	Not serious	N/A	Serious[1]	Serious[2]	144	MD -0.17 (-0.85, 0.51)	Low
BPSD: NPI (follow-up) – lower numbers favour intervention							
1 (Zhang 2015)	Not serious	N/A	Serious[1]	Serious[2]	144	MD -0.09 (-0.71, 0.53)	Low

1. Not a relevant intervention in the UK
2. Non-significant result
3. 95% CI crosses 1 line of a defined MID interval

Chinese Traditional medicine (Yokukansan) versus placebo

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
Cognition: MMSE (post-intervention) – higher numbers favour intervention							
1 (Farukawa 2017)	Serious[1]	N/A	Serious[2]	Serious[3]	137	MD -0.30 (-1.78, 1.18)	Very low
BPSD: NPI (post-intervention) – lower numbers favour intervention							
1 (Farukawa 2017)	Serious[1]	N/A	Serious[2]	Serious[3]	142	MD -0.40 (-1.84, 1.04)	Very low

1. No details of randomisation method or assessor blinding reported
2. Not a relevant intervention in the UK
3. Non-significant result

Chinese traditional medicine (Di-Huang-Yi-Zhi) and donepezil versus placebo and donepezil

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
Mini Mental State Examination – higher numbers favour Di-Huang-Yi-ZHI (@6 months)							

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
1 (Gu 2015)	Serious[1]	N/A	Serious[1]	Serious[2]	60	MD 0.85 (-0.72, 2.42)	Very low
Activities of Daily Living – lower numbers favour Di-Huang-Yi-ZHI (@6 months)							
1 (Gu 2015)	Very serious[4]	N/A	Serious[1]	Not serious	60	MD -6.54 (-9.84, -3.24)	Very low

1. No details of randomisation method or assessor blinding reported
2. Not a relevant intervention in the UK
3. Non-significant result
4. No details of randomisation method or assessor blinding reported; unclear what outcome measure used for ADL

Nutritional Formulation versus placebo

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
Neuropsychiatric Inventory – lower numbers favour nutritional formulation (@3 months)							
1 (Remington 2014)	Serious[1]	N/A	Not Serious	Serious[2]	83	MD 0.40 (-4.49, 5.29)	Low
Activities of Daily Living – lower numbers favour nutritional formulation (@3 months)							
1 (Remington 2014)	Serious[1]	N/A	Not Serious	Serious[2]	83	MD 2.30 (-5.51, 10.11)	Low

1. High number of participants lost to follow up
2. Non-significant result

Nutritional formulation consist of - 400µg folic acid, 6µg B1, 30I.U. alpha-tocopherol,400g SAM (200mg active ion), 600mg NAC and 500mg ALCAR

G.9.1.10 Music therapy

Music therapy versus standard care in people with dementia (post-intervention)

Full population

Quality assessment						No of participants		Effect estimate	Quality
No of publications	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Music therapy	Standard care	Summary of results Mean difference (95% CI)	
Cognition: MMSE – higher values favour intervention									

Quality assessment						No of participants		Effect estimate	Quality
No of publications	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Music therapy	Standard care	Summary of results Mean difference (95% CI)	
5	RCT	Serious[4]	Not serious	Serious[1]	Not serious	157	127	MD 1.91 (0.05, 3.78)	Low
Behavioural and psychological symptoms: NPI – lower values favour intervention									
1 (Raglio 2015)	RCT	Serious[4]	Not serious	N/A	Serious[2]	80	40	MD 0.72 (-4.38, 5.82)	Low
Depression: CSDD – lower values favour intervention									
1 (Chu 2014)	RCT	Serious[4]	Not serious	N/A	Not serious	49	51	MD -7.25 (-10.55, -3.95)	Moderate
Depression (standardised mean difference): CSDD or GDS – lower values favour intervention									
3	RCT	Serious[4]	Not serious	Serious[1]	Serious[5]	90	86	SMD -0.72 (-1.50, 0.05)	Very low
Agitation: CMAI – lower values favour intervention									
6	RCT	Serious[4]	Not serious	Serious[1]	Serious[2]	165	157	MD -4.67 (-9.67, 0.33)	Very low
Activities of daily living: Katz Index – higher values favour intervention									
1 (Ceccato 2012)	RCT	Serious[4]	Not serious	N/A	Very serious[2,3]	19	15	MD -0.67 (-1.20, -0.14)	Very low
HRQoL: QoL-AD – higher values favour intervention									
1 (Sarkamo 2016)	RCT	Serious[4]	Not serious	N/A	Serious[2]	51	23	MD 1.61 (-0.31, 3.53)	Low
HRQoL (standardised mean difference): QoL-AD or ADRQL or CBS– higher values favour intervention									
3	RCT	Serious[4]	Not serious	Not serious	Serious[5]	152	84	SMD 0.16 (-0.11, 0.43)	Low
Carer burden: ZBI – lower values favour intervention									
1 (Sarkamo 2016)	RCT	Serious[4]	Not serious	N/A	Serious[2]	51	23	MD -0.82 (-4.56, 2.92)	Low

Quality assessment						No of participants		Effect estimate	Quality
No of publications	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Music therapy	Standard care	Summary of results Mean difference (95% CI)	
Carer burden (standardised mean difference): ZBI or Global rating – lower values favour intervention									
2	RCT	Serious[4]	Not serious	Serious[1]	Serious[2]	77	36	SMD -0.40 (-0.91, 0.12)	Low

1. I^2>40%
2. Non-significant result
3. Low participant numbers
4. Issues with blinding of participants, personnel and/or assessor; personnel enthusiasm and training could influence outcome
5. 95% CI crosses 1 line of a defined MID interval

ADRQL: Alzheimer's Disease Related Quality of Life; CBS: Cornell Brown Scale for quality of life; CMAI: ; Cohen-Mansfield Agitation Inventory; CSDD: Cornell Scale of Depression in Dementia; GDS: Geriatric Depression Scale; HRQoL: health related quality of life; MMSE: Mini Mental State Examination; NPI: Neuropsychiatric inventory; QoL-AD: Quality of life in Alzheimer's disease; ZBI: Zarit Burden Interview

Sensitivity analysis excluding studies only recruiting people with non-cognitive symptoms (e.g. anxiety/depression) at baseline

Quality assessment						No of participants		Effect estimate	Quality
No of publications	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Music therapy	Standard care	Summary of results Mean difference (95% CI)	
Cognition: MMSE – higher values favour intervention									
5	RCT	Serious[4]	Not serious	Serious[1]	Not serious	157	127	MD 1.91 (0.05, 3.78)	Low
Depression: CSDD – lower values favour intervention									
1 (Chu 2014)	RCT	Serious[4]	Not serious	N/A	Not serious	49	51	MD -7.25 (-10.55, -3.95)	Moderate
Depression (standardised mean difference): CSDD or GDS – lower values favour intervention									
2	RCT	Serious[4]	Not serious	Serious[1]	Very serious[6]	76	74	SMD -0.40 (-1.18, 0.38)	Very low
Agitation: CMAI – lower values favour intervention									
2	RCT	Serious[4]	Not serious	Serious[1]	Serious[2]	165	157	MD -4.15 (-12.07, 3.76)	Very low
Activities of daily living: Katz Index – higher values favour intervention									

Quality assessment						No of participants		Effect estimate	Quality
No of publications	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Music therapy	Standard care	Summary of results Mean difference (95% CI)	
1 (Ceccato 2012)	RCT	Serious[4]	Not serious	N/A	Not serious	19	15	MD -0.67 (-1.20, -0.14)	Moderate
HRQoL: QoL-AD – higher values favour intervention									
1 (Sarkamo 2016)	RCT	Serious[4]	Not serious	N/A	Serious[2]	51	23	MD 1.61 (-0.31, 3.53)	Low
HRQoL (standardised mean difference): QoL-AD or ADRQL or CBS– higher values favour intervention									
1 (Sarkamo 2016)	RCT	Serious[4]	Not serious	Not serious	Serious[5]	51	23	SMD 0.35 (-0.14, 0.85)	Low
Carer burden: ZBI – lower values favour intervention									
1 (Sarkamo 2016)	RCT	Serious[4]	Not serious	N/A	Serious[2]	51	23	MD -0.82 (-4.56, 2.92)	Low

Quality assessment						No of participants		Effect estimate	Quality
No of publications	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Music therapy	Standard care	Summary of results Mean difference (95% CI)	
Carer burden (standardised mean difference): ZBI or Global rating – lower values favour intervention									
2	RCT	Serious[4]	Not serious	Serious[1]	Serious[2]	77	36	SMD -0.40 (-0.91, 0.12)	Low

1. I^2>40%
2. Non-significant result
3. Low participant numbers
4. Issues with blinding of participants, personnel and/or assessor; personnel enthusiasm and training could influence outcome
5. 95% CI crosses 1 line of a defined MID interval
6. 95% CI crosses 2 lines of a defined MID interval

ADRQL: Alzheimer's Disease Related Quality of Life; CBS: Cornell Brown Scale for quality of life; CMAI: ; Cohen-Mansfield Agitation Inventory; CSDD: Cornell Scale of Depression in Dementia; GDS: Geriatric Depression Scale; HRQoL: health related quality of life; MMSE: Mini Mental State Examination; NPI: Neuropsychiatric inventory; QoL-AD: Quality of life in Alzheimer's disease; ZBI: Zarit Burden Interview

Music therapy versus standard care in people with dementia (follow-up)

Full population

Quality assessment						No of participants		Effect estimate	Quality
No of publications	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Music therapy	Standard care	Summary of results Mean difference (95% CI)	
Cognition: MMSE – higher values favour intervention									
2	RCT	Serious[4]	Not serious	Not serious	Serious[1]	100	74	MD 1.53 (-0.27, 3.33)	Low
Behavioural and psychological symptoms: NPI – lower values favour intervention									
1 (Raglio 2015)	RCT	Serious[4]	Not serious	Not serious	Serious[1]	80	40	MD 1.90 (-3.71, 7.50)	Low
Depression: CSDD – lower values favour intervention									
1 (Chu 2014)	RCT	Serious[4]	Not serious	N/A	Serious[1]	49	51	MD -1.89 (-5.49, 1.71)	Low
Depression (standardised mean difference): CSDD or GDS– lower values favour intervention									

2	RCT	Serious[4]	Not serious	Serious[2]	Very serious[3]	62	62	SMD -0.61 (-1.57, 0.35)	Very low
Agitation: CMAI – lower values favour intervention									
2	RCT	Serious[4]	Not serious	Serious[2]	Not serious	66	68	MD -9.27 (-14.06, -4.48)	Low
HRQoL: QoL-AD – higher values favour intervention									
1 (Sarkamo 2016)	RCT	Serious[4]	Not serious	N/A	Not serious	51	23	MD 2.30 (0.01, 4.58)	Moderate
HRQoL (standardised mean difference): QoL-AD or CBS– higher values favour intervention									
2	RCT	Serious[4]	Not serious	Not serious	Serious[5]	152	84	SMD 0.35 (0.05, 0.65)	Low
Carer burden: ZBI – lower values favour intervention									
1 (Sarkamo 2016)	RCT	Serious[4]	Not serious	Not serious	Serious[1]	51	23	MD -1.74 (-5.83, 2.35)	Low
Carer burden (standardised mean difference): ZBI or Global rating – lower values favour intervention									
2	RCT	Serious[4]	Not serious	Serious[2]	Serious[5]	77	36	SMD -0.69 (-1.37, -0.01)	Very low

1. Non-significant result
2. I^2>40%
3. 95% CI crosses 2 lines of a defined MID interval
4. Issues with blinding of participants, personnel and/or assessor; personnel enthusiasm and training could influence outcome
5. 95% CI crosses 1 line of a defined MID interval

ADRQL: Alzheimer's Disease Related Quality of Life; CBS: Cornell Brown Scale for quality of life; CMAI: ; Cohen-Mansfield Agitation Inventory; CSDD: Cornell Scale of Depression in Dementia; GDS: Geriatric Depression Scale; HRQoL: health related quality of life; MMSE: Mini Mental State Examination; NPI: Neuropsychiatric inventory; QoL-AD: Quality of life in Alzheimer's disease; ZBI: Zarit Burden Interview

Sensitivity analysis excluding studies only recruiting people with non-cognitive symptoms (e.g. anxiety/depression) at baseline

Quality assessment						No of participants		Effect estimate	Quality
No of publications	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Music therapy	Standard care	Summary of results Mean difference (95% CI)	
Cognition: MMSE – higher values favour intervention									
2	RCT	Serious[4]	Not serious	Not serious	Serious[1]	100	74	MD 1.53 (-0.27, 3.33)	Low
Depression: CSDD – lower values favour intervention									
1 (Chu 2014)	RCT	Serious[4]	Not serious	N/A	Serious[1]	49	51	MD -1.89 (-5.49, 1.71)	Low

Depression (standardised mean difference): CSDD or GDS– lower values favour intervention									
1 (Chu 2014)	RCT	Serious[4]	Not serious	N/A	Very serious[3]	49	51	SMD -0.20 (-0.59, 0.20)	Very low
Agitation: CMAI – lower values favour intervention									
1 (Lin 2011)	RCT	Serious[4]	Not serious	N/A	Not serious	49	51	MD -7.40 (-11.26, -3.54)	Moderate
HRQoL: QoL-AD – higher values favour intervention									
1 (Sarkamo 2016)	RCT	Serious[4]	Not serious	N/A	Not serious	51	23	MD 2.30 (0.01, 4.58)	Moderate
HRQoL (standardised mean difference): QoL-AD or CBS– higher values favour intervention									
1 (Sarkamo 2016)	RCT	Serious[4]	Not serious	Not serious	Serious[5]	152	84	SMD 0.49 (-0.01, 0.99)	Low
Carer burden: ZBI – lower values favour intervention									
1 (Sarkamo 2016)	RCT	Serious[4]	Not serious	Not serious	Serious[1]	51	23	MD -1.74 (-5.83, 2.35)	Low
Carer burden (standardised mean difference): ZBI or Global rating – lower values favour intervention									
2	RCT	Serious[4]	Not serious	Serious[2]	Serious[5]	77	36	SMD -0.69 (-1.37, -0.01)	Very low

1. Non-significant result
2. I^2>40%
3. 95% CI crosses 2 lines of a defined MID interval
4. Issues with blinding of participants, personnel and/or assessor; personnel enthusiasm and training could influence outcome
5. 95% CI crosses 1 line of a defined MID interval

ADRQL: Alzheimer's Disease Related Quality of Life; CBS: Cornell Brown Scale for quality of life; CMAI: ; Cohen-Mansfield Agitation Inventory; CSDD: Cornell Scale of Depression in Dementia; GDS: Geriatric Depression Scale; HRQoL: health related quality of life; MMSE: Mini Mental State Examination; NPI: Neuropsychiatric inventory; QoL-AD: Quality of life in Alzheimer's disease; ZBI: Zarit Burden Interview

Music therapy versus active control in people with dementia (post-intervention)

Full population

Quality assessment						**No of participants**		**Effect estimate**	**Quality**
No of publications	**Design**	**Risk of bias**	**Indirectness**	**Inconsistency**	**Imprecision**	**Music therapy**	**Active comparator**	**Summary of results Mean difference (95% CI)**	
Cognition: MMSE – higher values favour intervention									

Quality assessment						No of participants		Effect estimate	Quality
No of publications	**Design**	**Risk of bias**	**Indirectness**	**Inconsistency**	**Imprecision**	**Music therapy**	**Active comparator**	**Summary of results Mean difference (95% CI)**	
1 (van der Winkel 2004)	RCT	Serious[4]	Not serious	N/A	Very serious[1,2]	15	11	MD 2.46 (-0.93, 5.85)	Very low
Cognition (standardised mean difference): MMSE or SIB – higher values favour intervention									
2	RCT	Serious[4]	Not serious	Not serious	Very serious[3]	33	30	SMD 0.23 (-0.27, 0.73)	Very low
Behavioural and psychological symptoms: NPI – lower values favour intervention									
1 (Narme 2014)	RCT	Serious[4]	Not serious	N/A	Very serious[1,2]	18	19	MD 1.20 (-6.67, 9.07)	Very low
Depression: GDS – lower values favour intervention									
1 (Cooke 2010)	RCT	Serious[4]	Not serious	N/A	Serious[1]	24	23	MD 0.23 (-0.31, 0.77)	Low
Agitation: CMAI – lower values favour intervention									
3	RCT	Serious[4]	Not serious	Not serious	Serious[1]	45	59	MD 2.82 (-1.61, 7.26)	Low
HRQoL: Dementia Quality of Life – higher values favour intervention									
1 (Cooke 2010)	RCT	Serious[4]	Not serious	N/A	Serious[1]	24	23	MD 0.09 (-1.47, 1.65)	Low

Quality assessment						No of participants		Effect estimate	Quality
No of publications	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Music therapy	Active comparator	Summary of results Mean difference (95% CI)	
Carer burden: NPI distress – lower values favour intervention									
1 (Narme 2014)	RCT	Serious[4]	Not serious	N/A	Very serious[1,2]	18	19	MD 0.90 (-2.40, 4.20)	Very low

1. Non-significant result
2. Low patient numbers
3. 95% CI crosses 2 lines of a defined MID interval
4. Issues with blinding of participants, personnel and/or assessor; personnel enthusiasm and training could influence outcome

CMAI: Cohen-Mansfield Agitation Inventory; MMSE: Mini Mental State Examination; NPI: Neuropsychiatric inventory; SIB: Severe Impairment Battery; ZBI: Zarit Burden Interview

Sensitivity analysis excluding studies only recruiting people with non-cognitive symptoms (e.g. anxiety/depression) at baseline

Quality assessment						No of participants		Effect estimate	Quality
No of publications	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Music therapy	Active comparator	Summary of results Mean difference (95% CI)	
Cognition (standardised mean difference): MMSE or SIB – higher values favour intervention									
1 (Narme 2014)	RCT	Serious[4]	Not serious	N/A	Very serious[3]	18	19	SMD 0.05 (-0.59, 0.70)	Very low
Behavioural and psychological symptoms: NPI – lower values favour intervention									
1 (Narme 2014)	RCT	Serious[4]	Not serious	N/A	Very serious[1,2]	18	19	MD 1.20 (-6.67, 9.07)	Very low
Agitation: CMAI – lower values favour intervention									
1 (Narme 2014)	RCT	Serious[4]	Not serious	N/A	Serious[1]	18	19	MD 5.90 (-2.08, 13.88)	Low

Quality assessment						No of participants		Effect estimate	Quality
No of publications	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Music therapy	Active comparator	Summary of results Mean difference (95% CI)	
Carer burden: NPI distress – lower values favour intervention									
1 (Narme 2014)	RCT	Serious[4]	Not serious	N/A	Very serious[1,2]	18	19	MD 0.90 (-2.40, 4.20)	Very low

1. Non-significant result
2. Low patient numbers
3. 95% CI crosses 2 lines of a defined MID interval
4. Issues with blinding of participants, personnel and/or assessor; personnel enthusiasm and training could influence outcome

CMAI: Cohen-Mansfield Agitation Inventory; MMSE: Mini Mental State Examination; NPI: Neuropsychiatric inventory; SIB: Severe Impairment Battery; ZBI: Zarit Burden Interview

Music therapy versus active control in people with dementia (follow-up)

Full population

Quality assessment						No of participants		Effect estimate	Quality
No of publications	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Music therapy	Active comparator	Summary of results Mean difference (95% CI)	
Cognition: SIB – higher values favour intervention									
1 (Narme 2014)	RCT	Serious[3]	Not serious	N/A	Very serious[1,2]	18	19	MD 0.90 (-10.77, 12.57)	Very low
Behavioural and psychological symptoms: NPI – lower values favour intervention									
1 (Narme 2014)	RCT	Serious[3]	Not serious	N/A	Very serious[1,2]	18	19	MD -2.10 (-10.51, 6.31)	Very low
Agitation: CMAI – lower values favour intervention									
2	RCT	Serious[3]	Not serious	Not serious	Serious[1]	35	53	MD 3.03 (-1.43, 7.49)	Low
Carer burden: ZBI – lower values favour intervention									
1 (Narme 2014)	RCT	Serious[3]	Not serious	N/A	Very serious[1,2]	18	19	MD -1.20 (-5.07, 2.67)	Very low

1. Non-significant result

Quality assessment						No of participants		Effect estimate	Quality
No of publications	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Music therapy	Active comparator	Summary of results Mean difference (95% CI)	

2. Low patient number
3. Issues with blinding of participants, personnel and/or assessor; personnel enthusiasm and training could influence outcome

MMSE: Mini Mental State Examination; NPI: Neuropsychiatric inventory; SIB: Severity Impairment Battery; ZBI: Zarit Burden Interview

Sensitivity analysis excluding studies only recruiting people with non-cognitive symptoms (e.g. anxiety/depression) at baseline

Quality assessment						No of participants		Effect estimate	Quality
No of publications	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Music therapy	Active comparator	Summary of results Mean difference (95% CI)	
Cognition: SIB – higher values favour intervention									
1 (Narme 2014)	RCT	Serious[3]	Not serious	N/A	Very serious[1,2]	18	19	MD 0.90 (-10.77, 12.57)	Very low
Behavioural and psychological symptoms: NPI – lower values favour intervention									
1 (Narme 2014)	RCT	Serious[3]	Not serious	N/A	Very serious[1,2]	18	19	MD -2.10 (-10.51, 6.31)	Very low
Agitation: CMAI – lower values favour intervention									
1 (Narme 2014)	RCT	Serious[3]	Not serious	N/A	Serious[1]	18	19	MD 6.40 (-1.49, 14.29)	Low
Carer burden: ZBI – lower values favour intervention									
1 (Narme 2014)	RCT	Serious[3]	Not serious	N/A	Very serious[1,2]	18	19	MD -1.20 (-5.07, 2.67)	Very low

1. Non-significant result
2. Low patient number
3. Issues with blinding of participants, personnel and/or assessor; personnel enthusiasm and training could influence outcome

MMSE: Mini Mental State Examination; NPI: Neuropsychiatric inventory; SIB: Severity Impairment Battery; ZBI: Zarit Burden Interview

G.9.1.11 Aromatherapy

Quality assessment						No of participants		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Aromatherapy	Control	Summary of results Mean difference (95% CI)	
Behavioural and psychological symptoms – lower values favour intervention									
Post-intervention – NPI									
1 (Burns 2011)	RCT	Serious	Not serious	N/A	Serious[1]	32	31	MD 2.80 (-6.15, 11.75)	Low
Agitation – lower values favour intervention									
Post-intervention (standardised mean difference) – CMAI or PAS									
3	RCTs	Serious	Not serious	Serious[2]	Very serious[3]	94	96	SMD -0.43 (-1.08, 0.23)	Very low
Post-intervention – CMAI									
2	RCT	Serious	Not serious	Serious[2]	Serious[1]	62	65	MD -9.36 (-22.01, 3.30)	Low
Depression – lower values favour intervention									
Post-intervention – CSDD									
1 (Yang 2016)	RCT	Serious	Not serious	N/A	Not serious	27	29	MD -5.83 (-8.57, -3.09)	Moderate
Activities of daily living – higher values favour intervention									
Post-intervention – Barthel Index									
1 (Burns 2011)	RCT	Serious	Not serious	N/A	Serious[1]	32	31	MD -0.50 (-1.81, 0.81)	Low
Quality of life – higher values favour intervention									
Post-intervention – Blau QoL									
1 (Burns 2011)	RCT	Serious	Not serious	N/A	Serious[1]	32	31	MD 19.00 (-24.87, 62.87)	Low

1. Non-significant result
2. $i^2 > 40\%$
3. 95% CI crosses 2 lines of a defined MID interval

CMAI: Cohen-Mansfield Agitation Inventory; CSDD: Cornell Scale for Depression in Dementia; MD: mean difference; NPI: Neuropsychiatric inventory; PAS: Pittsburgh agitation scale; QoL: Quality of life; RCT: randomised control trial; SMD: standardised mean difference

G.9.1.12 Light therapy in people with dementia

Full population

Quality assessment						No of participants		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Light therapy	Control	Summary of results Mean difference (95% CI)	
Cognition: MMSE – higher values favour intervention									
Post-intervention									
2	RCTs	Serious	Not serious	Not serious	Serious[1]	31	33	MD 0.68 (-2.46, 3.81)	Low
Follow-up									
1 (Burns 2009)	RCT	Serious	Not serious	N/A	Serious[1]	22	24	MD 0.00 (-3.21, 3.21)	Low
Behavioural and psychological symptoms: MOUSEPAD – lower values favour intervention									
Post-intervention									
1 (Burns 2009)	RCT	Serious	Not serious	N/A	Serious[1]	22	25	MD -0.10 (-3.81, 3.61)	Low
Follow-up									
1 (Burns 2009)	RCT	Serious	Not serious	N/A	Serious[1]	22	23	MD 0.20 (-3.39, 3.79)	Low
Depression: CSDD – lower values favour intervention									
Post-intervention									
2	RCTs	Serious	Not serious	Serious[2]	Serious[1]	51	52	MD -3.33 (-9.63, 2.98)	Very low
Follow-up									
1 (Burns 2009)	RCT	Serious	Not serious	N/A	Serious[1]	21	24	MD -0.20 (-1.85, 1.45)	Low
Agitation: CMAI – lower values favour intervention									
Post-intervention									
2	RCTs	Serious	Not serious	Serious[2]	Serious[1]	52	56	MD -12.32 (-28.76, 4.12)	Very low
Follow-up									
1 (Burns 2009)	RCT	Serious	Not serious	N/A	Serious[1]	22	24	MD -4.50 (-11.61, 2.61)	Low

Quality assessment						No of participants		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Light therapy	Control	Summary of results Mean difference (95% CI)	
Activities of daily living: CRBRS – higher values favour intervention									
Post-intervention									
1 (Burns 2009)	RCT	Serious	Not serious	N/A	Serious[1]	22	25	MD -0.10 (-1.43, 1.23)	Low
Follow-up									
1 (Burns 2009)	RCT	Serious	Not serious	N/A	Serious[1]	22	21	MD 1.00 (-0.78, 2.78)	Low

[1] Non-significant result
[2] I^2>40%
CMAI: Cohen-Mansfield Agitation Inventory; CRBRS: Crichton Royal Behavior Rating Scale; CSDD: Cornell Scale for Depression in Dementia; MMSE: Mini Mental State Examination; MOUSEPAD: Manchester and Oxford Universities Scale for the Psychological Assessment of Dementia

Sensitivity analysis excluding studies only recruiting people with non-cognitive symptoms (e.g. anxiety/depression) at baseline

Quality assessment						No of participants		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Light therapy	Control	Summary of results Mean difference (95% CI)	
Cognition: MMSE – higher values favour intervention									
Post-intervention									
1 (Graf 2001)	RCT	Very serious	Not serious	N/A	Serious[1]	9	9	MD 2.60 (-3.00, 8.20)	Low
Depression: CSDD – lower values favour intervention									
Post-intervention									
1 (Onega 2016)	RCT	Serious	Not serious	N/A	Not serious	30	30	MD -6.53 (-8.69, -4.37)	Moderate
Agitation: CMAI – lower values favour intervention									
Post-intervention									
1 (Onega 2016)	RCT	Serious	Not serious	N/A	Not serious	30	30	MD -20.39 (-29.57, -11.21)	Moderate

CMAI: Cohen-Mansfield Agitation Inventory; CSDD: Cornell Scale for Depression in Dementia; MMSE: Mini Mental State Examination

G.9.1.13 Non-invasive brain stimulation

Non-invasive brain stimulation in people with Alzheimer's disease (post-intervention)

Quality assessment						No of participants		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Brain stimulation	Sham	Summary of results Mean difference (95% CI)	
Cognition: MMSE – higher values favour intervention									
4	RCT	Serious[3]	Not serious	Not serious	Serious[1]	50	40	MD 0.79 (-0.57, 2.15)	Low
Cognition (standardised mean difference): MMSE or ADAS-cog – higher values favour intervention									
5	RCT	Serious[3]	Not serious	Not serious	Serious[1]	57	48	SMD 0.28 (-0.12, 0.68)	Low
Activities of daily living: IADL – higher values favour intervention									
2	RCT	Serious[3]	Not serious	Not serious	Serious[1,2]	17	16	MD 0.00 (-1.45, 1.45)	Low
Depression: Geriatric Depression Scale (GDS)– lower values favour intervention									
2	RCT	Serious[3]	Not serious	Not serious	Serious[1]	33	23	MD -1.08 (-2.24, 0.08)	Low

1. Non-significant result
2. Low participant numbers
3. No information on randomisation and allocation concealment methods and assessor blinding, unclear whether groups were balanced at baseline for some outcomes of interest

ADAS-cog: Alzheimer's Disease Assessment Scale-cognitive; IADL: Instrumental Activities of Daily Living; MMSE: Mini Mental State

Non-invasive brain stimulation in people with Alzheimer's disease (follow-up)

Quality assessment						No of participants		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Brain stimulation	Sham	Summary of results Mean difference (95% CI)	
Cognition: MMSE – higher values favour intervention									
3	RCT	Serious[4]	Not serious	Serious[1]	Serious[2]	45	35	MD 1.23 (-1.68, 4.14)	Very low
Activities of daily living: IADL – higher values favour intervention									

Quality assessment						No of participants		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Brain stimulation	Sham	Summary of results Mean difference (95% CI)	
1 (Cotelli 2014)	RCT	Serious[4]	Not serious	N/A	Very serious[2,3]	12	12	MD 0.10 (-1.58, 1.78)	Very low
Depression: GDS – lower values favour intervention									
2	RCT	Serious[4]	Not serious	Not serious	Serious[2]	33	23	MD -2.07 (-4.19, 0.05)	Low

1. I^2>40%
2. Non-significant result
3. Low participant numbers
4. No information on randomisation and allocation concealment methods and assessor blinding, unclear whether groups were balanced at baseline for some outcomes of interest

IADL: Instrumental Activities of Daily Living; GDS: Geriatric depression scale; MMSE: Mini Mental State Examinations

G.9.1.14 Non-invasive brain stimulation in people with mild vascular dementia (post-intervention)

Quality assessment						No of participants		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Brain stimulation	Sham	Summary of results Mean difference (95% CI)	
Cognition: ADAS-cog – lower values favour intervention									
1 (Andre 2016)	RCT	Serious[3]	Not serious	N/A	Very serious[1,2]	13	8	MD 1.10 (-14.25, 16.45)	Very low

1. Non-significant result
2. Low participant numbers
3. No information on randomisation and allocation concealment methods and assessor blinding, unclear whether groups were balanced at baseline for some outcomes of interest

ADAS-cog: Alzheimer's Disease Assessment Scale-cognitive

G.9.1.15 Acupuncture

Quality assessment						No of participants		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Acupuncture	No treatment	Summary of results Mean difference (95% CI)	
Cognition: MMSE – higher values favour intervention									
Post-intervention									
2	RCTs	Very serious[3]	Not serious	Serious[1]	Serious[2]	111	112	MD 1.88 (-3.31, 7.07)	Very low
Activities of daily living: Barthel Index – higher values favour intervention									
Post-intervention									
1 (Wang 2014)	RCT	Serious[4]	Not serious	N/A	Serious[2]	27	28	MD 1.60 (-0.94, 4.14)	Low

1. I^2>40%
2. Non-significant result
3. Unclear reporting of methods
4. Lack of blinding in study

MMSE: Mini Mental State Examination

G.9.1.16 Assisted animal therapy

Quality assessment						No of participants		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Assisted animal therapy	Control	Summary of results Mean difference (95% CI)	
Depression: CSDD (post-intervention) – lower values favour intervention									
1 (Olsen 2017)	RCT	Serious[1]	Not serious	Not serious	Serious[2]	22	25	MD -2.47 (-6.14, 1.21)	Low
Depression: CSDD (follow-up): Mild to moderate Dementia (CDR score 1 – 2) – lower values favour intervention									
1 (Olsen 2017)	RCT	Serious[1]	Not serious	N/A	Very serious[2,3]	11	14	MD -4.36 (-9.74, 1.02)	Very low
Depression: CSDD (follow-up): Severe Dementia (CDR score 3) – lower values favour intervention									
1 (Olsen 2017)	RCT	Serious[1]	Not serious	N/A	Not serious	11	10	MD -11.04 (-18.11, -3.97)	Moderate
Depression: CSDD (follow-up): All severities – lower values favour intervention									

Quality assessment						No of participants		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Assisted animal therapy	Control	Summary of results Mean difference (95% CI)	
1 (Olsen 2017)	RCT	Serious[1]	Not serious	Serious[4]	Not serious	22	24	MD -6.81 (-11.09, -2.53)	Low
Quality of life: QUALID (post-intervention) – lower values favour intervention									
1 (Olsen 2017)	RCT	Serious[1]	Not serious	Not serious	Serious[2]	24	26	SMD -0.14 (-0.70, 0.42)	Low
Quality of life: QUALID (follow-up): Mild to moderate Dementia (CDR score 1 – 2) – lower values favour intervention									
1 (Olsen 2017)	RCT	Serious[1]	Not serious	N/A	Very serious[2, 3]	12	14	SMD -0.24 (-0.53, 1.02)	Very low
Quality of life: QUALID (follow-up): Severe Dementia (CDR score 3) – lower values favour intervention									
1 (Olsen 2017)	RCT	Serious[1]	Not serious	N/A	Not serious	11	11	SMD -0.91 (-1.80, -0.02)	Moderate
Quality of life: QUALID (follow-up): lower values favour intervention									
1 (Olsen 2017)	RCT	Serious[1]	Not serious	Serious[4]	Serious[2]	23	25	SMD -0.26 (-0.84, 0.33)	Very low

1. Method of diagnosis of dementia is not reported.
2. Non-significant result.
3. Low participant numbers.
4. I^2>40%

Note: data required for analysis was calculated by information provided in Olsen 201, but not reported in Olsen 2017.
BARS: Brief Agitation Rating Scale, CSDD: Cornell Scale for Depression in Dementia; QUALID: Quality of Life in Late-stage Dementia

G.9.1.17 Robotic pet therapy

Quality assessment						No of participants		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Robotic pet therapy	Usual care	Summary of results Mean difference (95% CI)	
Depression: CSDD (post-intervention) – lower values favour intervention									
1 (Petersen 2017)	RCT	Not serious	Not serious	N/A	Not serious	35	26	MD -2.03 (-1.83, -2.23)	High

CSDD: Cornell Scale for Depression in Dementia, RAID: Rating for Anxiety in Dementia.

G.9.1.18 Adapted mindfulness program

Quality assessment						No of participants		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Adapted mindfulness	Usual care	Summary of results Mean difference (95% CI)	
Cognition (MMSE) higher values favour intervention									
1 Churcher Clarke (2017)	RCT	Very serious[1]	Not serious	N/A	Serious[2]	20	8	MD 1.65 (-2.52, 5.82)	Very low
Quality of life (QOLAD) higher values favour intervention									
1 Churcher Clarke (2017)	RCT	Very serious[1]	Not serious	N/A	Not serious	20	8	MD 4.14 (0.46, 7.82)	Low
Depression (CSDD) lower values favour intervention									
1 Churcher Clarke (2017)	RCT	Very serious[1]	Not serious	N/A	Serious[2]	20	8	MD 1.58 (-3.12, 6.28)	Very low

1. Single blind, limited reporting pilot study
2. Non-significant result

G.9.1.19 Home safety toolkit

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Usual care	Summary of results	
Revised Scale for Caregiving Self-efficacy (higher numbers favour intervention)									
1 (Horvath 2013)	RCT	Not serious	Not serious	N/A	Serious[1]	60	48	MD 44.65 (-31.50, 120.80)	Moderate
MBRC Caregiver Strain Instrument (lower numbers favour intervention)									
1 (Horvath 2013)	RCT	Not serious	Not serious	N/A	Serious[1]	60	48	MD -1.01 (-2.36, 0.34)	Moderate
Home Safety Checklist (lower numbers favour intervention)									
1 (Horvath 2013)	RCT	Not serious	Not serious	N/A	Serious[1]	60	48	MD -4.26 (-11.89, 3.37)	Moderate
Risky Behaviour Questionnaire (lower numbers favour intervention)									
1 (Horvath 2013)	RCT	Not serious	Not serious	N/A	Serious[1]	60	48	MD -3.49 (-16.82, 9.84)	Moderate

1. Non-significant result

G.9.2 Pre, peri and post-diagnostic counselling and support for people living with dementia and their families

- How effective are pre, peri & post-diagnostic counselling and support on outcomes for people living with dementia and their families?

G.9.2.1 Psychosocial interventions (outcomes in people with dementia)

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Control	Summary of results	
Quality of life (QoL-VAS) at 12 months – higher numbers favour intervention									
1 (Waldorff 2012)	RCT	Not serious[1]	Not serious	N/A	Serious[2]	128	143	MD 2.95 (-1.80, 7.70)	Moderate
Quality of life (QoL-VAS) at 36 months – higher numbers favour intervention									
2 (Koivisto 2016, Phung 2013)	RCT	Not serious[1]	Not serious	Not serious	Serious[2]	247	319	MD -2.18 (-7.11, 2.75)	Moderate
Quality of life (QoL-AD) at 12 months – higher numbers favour intervention									
1 (Waldorff 2012)	RCT	Not serious[1]	Not serious	N/A	Not serious	130	144	MD 2.14 (0.84, 3.44)	High
Quality of life (QoL-AD) at 36 months – higher numbers favour intervention									
2 (Koivisto 2016, Phung 2013)	RCT	Not serious[1]	Not serious	Not serious	Serious[2]	247	319	MD -0.62 (-1.91, 0.67)	Moderate
Cognitive impairment (MMSE) at 12 months – higher numbers favour intervention									
1 (Waldorff 2012)	RCT	Not serious[1]	Not serious	N/A	Serious[2]	130	139	MD 0.25 (-0.73, 1.23)	Moderate
Cognitive impairment (MMSE) at 36 months – higher numbers favour intervention									
2 (Koivisto 2016, Phung 2013)	RCT	Not serious[1]	Not serious	Not serious	Serious[2]	247	319	MD -0.40 (-1.73, 0.93)	Moderate
Memory disorder severity (CDR-SOB) at 36 months – lower numbers favours intervention									
1 (Koivisto 2016)	RCT	Not serious[1]	Not serious	N/A	Not serious	84	152	MD 1.30 (0.07, 2.53)	High
Activities of daily living (ADSC-ADL) at 12 months – higher numbers favour intervention									

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Control	Summary of results	
1 (Waldorff 2012)	RCT	Not serious[1]	Not serious	N/A	Serious[2]	130	143	MD -1.76 (-4.85, 1.33)	Moderate
Activities of daily living (ADSC-ADL) at 36 months – higher numbers favour intervention									
2 (Koivisto 2016, Phung 2013)	RCT	Not serious[1]	Not serious	Not serious	Not serious	247	319	MD -5.60 (-9.68, -1.53)	High
Behavioural disturbances (NPI-Q) at 12 months – lower numbers favour intervention									
1 (Waldorff 2012)	RCT	Not serious[1]	Not serious	N/A	Serious[2]	129	143	MD 0.42 (-0.55, 1.39)	Moderate
Behavioural disturbances (NPI or NPI-Q) at 36 months – lower numbers favour intervention									
2 (Koivisto 2016, Phung 2013)	RCT	Not serious[1]	Not serious	Not serious	Serious[2]	247	319	MD 0.34 (-0.93, 1.60)	Moderate
Depression (CDS) at 12 months – lower numbers favour intervention									
1 (Waldorff 2012)	RCT	Not serious[1]	Not serious	N/A	Not serious	130	141	MD -1.58 (-2.79, -0.37)	High
Depression (CDS) at 36 months – lower numbers favour intervention									
1 (Phung 2013)	RCT	Not serious[1]	Not serious	N/A	Serious[2]	163	167	MD -0.05 (-1.41, 1.31)	Moderate
Nursing home placement at 36 months – lower numbers favour intervention									
2 (Koivisto 2016, Phung 2013)	RCT	Not serious[1]	Not serious	Not serious	Serious[2]	247	319	RR 1.03 (0.77, 1.39)	Moderate
Mortality at 12 months – lower numbers favour intervention									
1 (Waldorff 2012)	RCT	Not serious[1]	Not serious	N/A	Serious[2]	163	167	RR 3.42 (0.96, 12.19)	Moderate
Mortality at 36 months – lower numbers favour intervention									
2 (Koivisto 2016, Phung 2013)	RCT	Not serious[1]	Not serious	Serious[3]	Serious[2]	247	319	RR 1.37 (0.69, 2.73)	Low

1. Participants in studies not blinded, but not judged to be a serious risk of bias
2. Non-significant result

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Control	Summary of results	
3. i^2>40%									
Waldorff 2012 and Phung 2013 report the 12-month and 36-month follow-up of the same RCT.									

G.9.2.2 Psychosocial interventions (outcomes in caregivers)

Quality assessment						No of caregivers		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Control	Summary of results	
Quality of life (QoL-VAS) at 12 months – higher numbers favour intervention									
1 (Waldorff 2012)	RCT	Not serious[1]	Not serious	N/A	Serious[2]	128	144	MD -0.51 (-4.46, 3.44)	Moderate
Quality of life (QoL-VAS) at 36 months – higher numbers favour intervention									
2 (Koivisto 2016, Phung 2013)	RCT	Not serious[1]	Not serious	Serious[3]	Serious[2]	247	319	MD 0.25 (-5.81, 6.30)	Low
Quality of life (QoL-15D) at 36 months – higher numbers favour intervention									
1 (Koivisto 2016)	RCT	Not serious[1]	Not serious	N/A	Serious[2]	84	152	MD 0.00 (-0.04, 0.03)	Moderate
Psychological distress during caregiving (GHQ) at 36 months – lower numbers favour intervention									
1 (Koivisto 2016)	RCT	Not serious[1]	Not serious	N/A	Serious[2]	84	152	MD -0.92 (-2.51, 0.67)	Moderate
Orientation to life (SOC) at 36 months – higher numbers favour intervention									
1 (Koivisto 2016)	RCT	Not serious[1]	Not serious	N/A	Serious[2]	84	152	MD 1.53 (-5.71, 8.77)	Moderate
Depression (GDS) at 12 months – lower numbers favour intervention									
1 (Waldorff 2012)	RCT	Not serious[1]	Not serious	N/A	Serious[2]	128	143	MD 0.70 (-0.47, 1.87)	Moderate
Depression (BDI or GDS) at 36 months – lower numbers favour intervention									
2 (Koivisto 2016, Phung 2013)	RCT	Not serious[1]	Not serious	N/A	Serious[2]	247	319	MD 0.07 (-1.85, 1.99)	Moderate

1. Participants in studies not blinded, but not judged to be a serious risk of bias
2. Non-significant result

Quality assessment						No of caregivers		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Control	Summary of results	
3. i^2>40% Waldorff 2012 and Phung 2013 report the 12-month and 36-month follow-up of the same RCT.									

G.9.2.3 Self-management interventions (outcomes in people with dementia)

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Usual care	Summary of results	
Health-related quality of life (15D) at 9 months – higher favour intervention									
1 (Laakkonen 2016)	RCT	Serious[1]	Not serious	N/A	Serious[2]	67	67	MD 0.01 (-0.02, 0.04)	Low
Global assessment (CDR) at 9 months – higher favour intervention									
1 (Laakkonen 2016)	RCT	Serious[1]	Not serious	N/A	Serious[2]	67	67	MD 0.53 (-0.09, 1.15)*	Low
Cognitive function (VF) at 9 months – higher favour intervention									
1 (Laakkonen 2016)	RCT	Serious[1]	Not serious	N/A	Not serious	67	67	MD 1.22 (0.31, 2.13)	Moderate
Cognitive function (CDT) at 9 months – higher favour intervention									
1 (Laakkonen 2016)	RCT	Serious[1]	Not serious	N/A	Not serious	67	67	MD 0.54 (0.05, 1.03)	Moderate
1. There was no blinding; baseline characteristics were not balanced between groups; control group received more than usual care; not all outcomes were reported 2. Non-significant result *Results were multiplied by -1 so direction of effect consistent with other cognitive outcomes to be included in a subgroup meta-analysis									

G.9.2.4 Self-management interventions (outcomes in spouses)

Quality assessment						No of caregivers		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Usual care	Summary of results	
Health-related quality of life (RAND-36 PCS) at 9 months – higher favour intervention									
1 (Laakkonen 2016)	RCT	Serious[1]	Not serious	N/A	Serious[2]	67	67	MD 1.70 (-0.31, 3.71)	Low
1. There was no blinding; baseline characteristics were not balanced between groups; control group received more than usual care; not all outcomes were reported 2. Non-significant result									

G.10 Managing non-cognitive symptoms

G.10.1 Interventions for treating illness emergent non-cognitive symptoms in people living with dementia

- What are the most effective pharmacological interventions for managing illness emergent non-cognitive symptoms, such as psychosis, depression, behavioural changes in people living with dementia?
- What are the most effective non-pharmacological interventions for managing illness emergent non-cognitive symptoms, such as psychosis, depression, behavioural changes in people living with dementia?

G.10.1.1 Anxiety and depression

Sertraline vs placebo (12-13 weeks)

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
Depression (Cornell Scale) – lower numbers favour sertraline							
3 (Banerjee, Lyketos, Weintraub)	Not serious	Serious[2]	Not serious	Serious[3]	348	MD -1.12 (-4.26, 2.01)	Low
Hamilton Depression Rating Scale – lower numbers favour sertraline							
1 (Lyketos)	Not serious	N/A	Not serious	Serious[3]	44	MD -4.10 (-8.77, 0.57)	Low
Improvement in mADCS-CGIC – higher numbers favour sertraline							
1 (Weintraub)	Not serious	N/A	Not serious	Serious[3]	131	OR 1.01 (0.52, 1.97)	Moderate
Mini Mental State Examination – higher numbers favour sertraline							
2 (Banerjee, Lyketos)	Not serious	Not serious	Not serious	Serious[3]	217	MD -0.25 (-1.48, 0.97)	Moderate
Activities of daily living – lower numbers favour sertraline							
2 (Banerjee, Lyketos)	Not serious	Serious[2]	Not serious	Serious[3]	217	SMD 0.10 (-0.46, 0.65)	Low
NPI – lower numbers favour sertraline							
2 (Banerjee, Lyketos)	Not serious	Not serious	Not serious	Serious[3]	217	MD 1.35 (-2.88, 5.58)	Moderate

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
Quality of life (patient-reported DEMQoL) – higher numbers favour sertraline							
1 (Banerjee)	Not serious	N/A	Not serious	Serious[3]	173	MD 0.30 (-3.40, 4.01)	Moderate
Quality of life (carer-reported DEMQoL) – higher numbers favour sertraline							
1 (Banerjee)	Serious[1]	N/A	Not serious	Serious[3]	173	MD -1.98 (-6.16, 2.21)	Low
Quality of life (patient-reported EQ-5D) – higher numbers favour sertraline							
1 (Banerjee)	Not serious	N/A	Not serious	Serious[3]	173	MD -3.44 (-10.86, 3.98)	Moderate
Quality of life (carer-reported EQ-5D) – higher numbers favour sertraline							
1 (Banerjee)	Serious[1]	N/A	Not serious	Serious[3]	173	MD 0.61 (-5.8, 6.59)	Low
Carer burden (Zarit) – lower numbers favour sertraline							
1 (Banerjee)	Not serious	N/A	Not serious	Serious[3]	173	MD -0.50 (-4.28, 3.27)	Moderate
Carer mental health (GHQ) – lower numbers favour sertraline							
1 (Banerjee)	Not serious	Not serious	Not serious	Not serious	173	MD 1.47 (0.06, 2.89)	High
SF-12 (physical) – higher numbers favour sertraline							
1 (Banerjee)	Not serious	Not serious	Not serious	Serious[3]	173	MD 1.28 (-1.48, 4.03)	Moderate
SF-12 (mental) – higher numbers favour sertraline							
1 (Banerjee)	Not serious	Not serious	Not serious	Not serious	173	MD -2.99 (-5.87, -0.11)	High

1. Proxy-reported outcomes.
2. i^2 value > 40%.
3. Non-significant result.

Sertraline vs placebo (24-39 weeks)

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
Depression (Cornell Scale) – lower numbers favour sertraline							
2 (Banerjee, Weintraub)	Not serious	Not serious	Not serious	Serious[3]	281	MD 0.16 (-1.16, 1.49)	Low
Improvement in mADCS-CGIC – higher numbers favour sertraline							
1 (Weintraub)	Not serious	N/A	Not serious	Serious[3]	131	OR 1.23 (0.64, 2.35)	Moderate
Mini Mental State Examination – higher numbers favour sertraline							

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
1 (Banerjee)	Not serious	N/A	Not serious	Serious[3]	150	MD -0.55 (-1.89, 0.79)	Moderate
Bristol Activities of Daily Living – lower numbers favour sertraline							
1 (Banerjee)	Not serious	N/A	Not serious	Serious[3]	150	MD 1.63 (-1.01, 4.27)	Moderate
NPI – lower numbers favour sertraline							
1 (Banerjee)	Not serious	N/A	Not serious	Serious[3]	150	MD 2.02 (-294, 6.97)	Moderate
Quality of life (patient-reported DEMQoL) – higher numbers favour sertraline							
1 (Banerjee)	Not serious	N/A	Not serious	Serious[3]	150	MD -1.76 (-5.75, 2.23)	Moderate
Quality of life (carer-reported DEMQoL) – higher numbers favour sertraline							
1 (Banerjee)	Serious[1]	N/A	Not serious	Serious[3]	150	MD 2.69 (-1.77, 7.15)	Low
Quality of life (patient-reported EQ-5D) – higher numbers favour sertraline							
1 (Banerjee)	Not serious	N/A	Not serious	Serious[3]	150	MD -4.34 (-12.56, 3.88)	Moderate
Quality of life (carer-reported EQ-5D) – higher numbers favour sertraline							
1 (Banerjee)	Serious[1]	N/A	Not serious	Serious[3]	150	MD -0.27 (-6.77, 6.24)	Low
Carer burden (Zarit) – lower numbers favour sertraline							
1 (Banerjee)	Not serious	N/A	Not serious	Serious[3]	150	MD -0.09 (-4.15, 3.98)	Moderate
Carer mental health (GHQ) – lower numbers favour sertraline							
1 (Banerjee)	Not serious	N/A	Not serious	Serious[3]	150	MD 0.43 (-1.09, 1.95)	Moderate
SF-12 (physical) – higher numbers favour sertraline							
1 (Banerjee)	Not serious	N/A	Not serious	Serious[3]	150	MD -1.68 (-4.58, 1.22)	Moderate
SF-12 (mental) – higher numbers favour sertraline							
1 (Banerjee)	Not serious	N/A	Not serious	Serious[3]	150	MD 0.09 (-2.94, 3.11)	Moderate
Any adverse events – lower numbers favour sertraline							
3 (Banerjee, Lyketos, Weintraub)	Not serious	Not serious	Not serious	Serious[4]	385	RR 1.59 (1.24, 2.05)	Moderate
Serious adverse events – lower numbers favour sertraline							
2 (Banerjee, Weintraub)	Not serious	Serious[2]	Not serious	Very serious[5]	347	RR 1.34 (0.51, 3.54)	Very low

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
1. Proxy-reported outcomes. 2. i^2 value > 40%. 3. Non-significant result. 4. 95% CI crosses one line of a defined MID interval. 5. 95% CI crosses two line of a defined MID interval.							

Mirtazapine vs placebo (13 weeks)

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
Depression (Cornell Scale) – lower numbers favour sertraline							
1 (Banerjee)	Not serious	N/A	Not serious	Serious[2]	180	MD 0.01 (-1.37, 1.38)	Moderate
Mini Mental State Examination – higher numbers favour sertraline							
1 (Banerjee)	Not serious	N/A	Not serious	Serious[2]	180	MD -0.27 (-1.48, 0.94)	Moderate
Bristol Activities of Daily Living – lower numbers favour sertraline							
1 (Banerjee)	Not serious	N/A	Not serious	Serious[2]	180	MD -0.04 (-2.44, 2.36)	Moderate
NPI – lower numbers favour sertraline							
1 (Banerjee)	Not serious	N/A	Not serious	Serious[2]	180	MD -3.56 (-8.07, 0.96)	Moderate
Quality of life (patient-reported DEMQoL) – higher numbers favour sertraline							
1 (Banerjee)	Not serious	N/A	Not serious	Serious[2]	180	MD -0.06 (-3.52, 3.39)	Moderate
Quality of life (carer-reported DEMQoL) – higher numbers favour sertraline							
1 (Banerjee)	Serious[1]	N/A	Not serious	Serious[2]	180	MD 3.13 (-1.09, 7.35)	Low
Quality of life (patient-reported EQ-5D) – higher numbers favour sertraline							
1 (Banerjee)	Not serious	N/A	Not serious	Serious[2]	180	MD 2.00 (-5.18, 9.19)	Moderate
Quality of life (carer-reported EQ-5D) – higher numbers favour sertraline							
1 (Banerjee)	Serious[1]	N/A	Not serious	Serious[2]	180	MD 3.62 (-2.31, 9.55)	Low
Carer burden (Zarit) – lower numbers favour sertraline							
1 (Banerjee)	Not serious	N/A	Not serious	Serious[2]	180	MD -1.11 (-4.93, 0.65)	Moderate
Carer mental health (GHQ) – lower numbers favour sertraline							
1 (Banerjee)	Not serious	N/A	Not serious	Serious[2]	180	MD -0.57 (-0.84, 1.98)	Moderate

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
SF-12 (physical) – higher numbers favour sertraline							
1 (Banerjee)	Not serious	N/A	Not serious	Serious[2]	180	MD -0.53 (-2.20, 3.26)	Moderate
SF-12 (mental) – higher numbers favour sertraline							
1 (Banerjee)	Not serious	N/A	Not serious	Serious[2]	180	MD 0.52 (-2.31, 3.36)	Moderate

1. Proxy-reported outcomes.
2. Non-significant result.

Mirtazapine vs placebo (39 weeks)

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
Depression (Cornell Scale) – lower numbers favour sertraline							
1 (Banerjee)	Not serious	N/A	Not serious	Serious[2]	158	MD -0.66 (-2.12, 0.79)	Moderate
Mini Mental State Examination – higher numbers favour sertraline							
1 (Banerjee)	Not serious	N/A	Not serious	Serious[2]	158	MD -1.71 (-2.48, 0.14)	Moderate
Bristol Activities of Daily Living – lower numbers favour sertraline							
1 (Banerjee)	Not serious	N/A	Not serious	Serious[2]	158	MD 1.19 (-1.37, 3.75)	Moderate
NPI – lower numbers favour sertraline							
1 (Banerjee)	Not serious	N/A	Not serious	Serious[2]	158	MD -1.51 (-6.25, 3.24)	Moderate
Quality of life (patient-reported DEMQoL) – higher numbers favour sertraline							
1 (Banerjee)	Not serious	N/A	Not serious	Serious[2]	158	MD -0.03 (-3.80, 3.75)	Moderate
Quality of life (carer-reported DEMQoL) – higher numbers favour sertraline							
1 (Banerjee)	Serious[1]	N/A	Not serious	Serious[2]	158	MD 3.69 (-0.77, 8.16)	Low
Quality of life (patient-reported EQ-5D) – higher numbers favour sertraline							
1 (Banerjee)	Not serious	N/A	Not serious	Serious[2]	158	MD -1.18 (-9.25, 6.89)	Moderate
Quality of life (carer-reported EQ-5D) – higher numbers favour sertraline							
1 (Banerjee)	Serious[1]	N/A	Not serious	Serious[2]	158	MD 1.11 (-7.44, 5.21)	Low
Carer burden (Zarit) – lower numbers favour sertraline							
1 (Banerjee)	Not serious	N/A	Not serious	Serious[2]	158	MD -2.80 (-6.99, 1.38)	Moderate

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
Carer mental health (GHQ) – lower numbers favour sertraline							
1 (Banerjee)	Not serious	N/A	Not serious	Serious[2]	158	MD -0.61 (-2.12, 0.90)	Moderate
SF-12 (physical) – higher numbers favour sertraline							
1 (Banerjee)	Not serious	N/A	Not serious	Serious[2]	158	MD 0.02 (-2.84, 2.88)	Moderate
SF-12 (mental) – higher numbers favour sertraline							
1 (Banerjee)	Not serious	N/A	Not serious	Serious[2]	158	MD -0.31 (-3.28, 2.66)	Moderate
Any adverse events – lower numbers favour sertraline							
1 (Banerjee)	Not serious	N/A	Not serious	Serious[3]	215	RR 1.56 (1.06, 2.30)	Moderate
Serious adverse events – lower numbers favour sertraline							
1 (Banerjee)	Not serious	N/A	Not serious	Very serious[4]	215	RR 0.92 (0.47, 1.82)	Low

1. Proxy-reported outcomes.
2. Non-significant result.
3. 95% CI crosses one line of a defined MID interval.
4. 95% CI crosses two line of a defined MID interval.

Psychological treatment vs usual care

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
Depression – lower numbers favour treatment							
6 (Ortega systematic review)	Serious[1]	Not serious	Not serious	Serious[4]	439	SMD -0.22 (-0.41, -0.03)	Low
Anxiety (RAID) – lower numbers favour treatment							
2 (Ortega systematic review)	Serious[1]	Not serious	Not serious	Not serious	65	MD -4.57 (-7.81, -1.32)	Moderate
Anxiety (self-rating) – lower numbers favour treatment							
2 (Ortega systematic review)	Serious[1]	Not serious	Not serious	Very serious[5]	65	SMD 0.05 (-0.44, 0.54)	Very low
Anxiety (NPI-A) – lower numbers favour treatment							

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
1 (Ortega systematic review)	Serious[1]	N/A	Not serious	Serious[3]	26	MD -2.40 (-4.96, 0.16)	Low
Quality of life (self-rating) – higher numbers favour treatment							
3 (Ortega systematic review)	Serious[1]	Not serious	Not serious	Serious[3]	334	MD 0.37 (-1.01, 1.75)	Low
Quality of life (proxy-rating) – higher numbers favour treatment							
2 (Ortega systematic review)	Serious[1]	Not serious	Not serious	Serious[3]	313	MD 0.66 (-0.77, 2.09)	Low
Activities of daily living – lower numbers favour treatment							
2 (Ortega systematic review)	Serious[1]	Not serious	Not serious	Serious[4]	313	SMD -0.13 (-0.35, 0.09)	Low
Neuropsychiatric symptoms – lower numbers favour treatment							
2 (Ortega systematic review)	Serious[1]	Serious[2]	Not serious	Very serious[5]	311	SMD -0.10 (-0.68, 0.48)	Very low
Mini Mental State Examination – higher numbers favour treatment							
4 (Ortega systematic review)	Serious[1]	Not serious	Not serious	Serious[3]	381	MD -0.97 (-2.01, 0.08)	Low
Caregiver depression – lower numbers favour treatment							
3 (Ortega systematic review)	Serious[1]	Serious[2]	Not serious	Very serious[5]	337	SMD -0.07 (-0.55, 0.41)	Very low

1. Lack of clarity about allocation concealment and blinding.
2. i^2 value > 40%.
3. Non-significant result.
4. 95% CI crosses one line of a defined MID interval.
5. 95% CI crosses two line of a defined MID interval.

PATH (Problem Adaptation Therapy) vs ST-CI (Supportive Therapy for Cognitively Impaired Older Adults)

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
Depression (MADRS) – lower numbers favour PATH							
1 (Kiosses)	Not serious	N/A	Serious[1]	Not serious	74	MD -0.60 (-1.06, -0.13)	Moderate
Depression (Rate of full remission: MADRS ≤7) – higher numbers favour PATH							
1 (Kiosses)	Not serious	N/A	Serious[1]	Serious[2]	74	HR 3.67 (1.20, 11.26)	Low
Depression (Rate of partial remission: MADRS ≤10) – higher numbers favour PATH							
1 (Kiosses)	Not serious	N/A	Serious[1]	Serious[2]	74	HR 2.85 (1.03, 7.91)	Low
Disability (WHODAS II) – lower numbers favour PATH							
1 (Kiosses)	Not serious	N/A	Serious[1]	Not serious	74	MD -0.67 (-1.14, -0.20)	Moderate

1. Study also contains people with mild cognitive impairment
2. 95% CI crosses one line of a defined MID interval

Structured depression management vs usual care (nursing-homes)

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
Depression prevalence (Cornell scale >7) – lower numbers favour intervention							
1 (Leontjevas)	Not serious	N/A	Not serious	Serious[1]	393	MD 0.6% (-5.6, 6.8)	Moderate
Depression prevalence (GDS8 >2) – lower numbers favour intervention							
1 (Leontjevas)	Not serious	N/A	Not serious	Serious[1]	393	MD -4.5% (-15.0, 6.0)	Moderate
Severe depression prevalence (Cornell scale >11) – lower numbers favour intervention							
1 (Leontjevas)	Not serious	N/A	Not serious	Serious[1]	393	MD 2.4% (-2.4, 7.2)	Moderate
Severe depression prevalence (GDS8 >4) – lower numbers favour intervention							
1 (Leontjevas)	Not serious	N/A	Not serious	Serious[1]	393	MD -0.3% (-0.8, 0.1)	Moderate
Depression (Cornell Scale) – lower numbers favour intervention							
1 (Leontjevas)	Not serious	N/A	Not serious	Serious[1]	393	MD 0.3 (-0.3, 0.9)	Moderate
Depression (GDS8) – lower numbers favour intervention							
1 (Leontjevas)	Not serious	N/A	Not serious	Serious[1]	393	MD -0.3 (-0.7, 0.1)	Moderate
EQ-VAS – higher numbers favour intervention							

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
1 (Leontjevas)	Not serious	N/A	Not serious	Not serious	393	MD 3.4 (0.5, 6.3)	High
1. Non-significant result.							

Psychogeriatric management vs usual care

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
Depression z score* – lower numbers favour psychogeriatric case management							
1 (Brodaty)	Not serious	N/A	Not serious	Serious[1]	44	MD 0.03 (-0.65, 0.72)	Moderate
Depression z score* – lower numbers favour psychogeriatric consultation							
1 (Brodaty)	Not serious	N/A	Not serious	Serious[1]	45	MD -0.11 (-0.95, 0.74)	Moderate
Psychosis z score* – lower numbers favour psychogeriatric case management							
1 (Brodaty)	Not serious	N/A	Not serious	Serious[1]	393	MD 0.31 (-0.42, 1.04)	Moderate
Psychosis z score* – lower numbers favour psychogeriatric consultation							
1 (Brodaty)	Not serious	N/A	Not serious	Serious[1]	393	MD 0.25 (-0.50, 1.00)	Moderate
*Calculated as the highest standardised score on any of the trial outcome measures for that individual 1. Non-significant result.							

Ambient bright light vs standard lighting

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
Depression in men with bright morning light (Cornell Scale) – lower numbers favour intervention							
1 (Hickman)	Very serious[1]	N/A	Not serious	Not serious	66	MD 2.62 (0.72, 4.52)	Low
Depression in men with bright evening light (Cornell Scale) – lower numbers favour intervention							
1 (Hickman)	Very serious[1]	N/A	Not serious	Serious[2]	66	MD 1.13 (-0.69, 2.95)	Very low
Depression in men with bright all-day light (Cornell Scale) – lower numbers favour intervention							
1 (Hickman)	Very serious[1]	N/A	Not serious	Serious[2]	66	MD 1.64 (-0.20, 3.48)	Very low
Depression in women with bright morning light (Cornell Scale) – lower numbers favour intervention							
1 (Hickman)	Very serious[1]	N/A	Not serious	Serious[2]	66	MD -1.61 (-3.49, 0.27)	Very low
Depression in women with bright evening light (Cornell Scale) – lower numbers favour intervention							

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
1 (Hickman)	Very serious[1]	N/A	Not serious	Serious[2]	66	MD 0.09 (-2.11, 2.29)	Very low
Depression in women with bright all-day light (Cornell Scale) – lower numbers favour intervention							
1 (Hickman)	Very serious[1]	N/A	Not serious	Serious[2]	66	MD 1.41 (-0.55, 3.37)	Very low

1. Crossover design with potentially serious confounding. Outcome assessment not adequately blinded.
2. Non-significant result.

Active music therapy vs reading

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
Quality of life (DQOL) – higher numbers favour intervention							
1 (Cooke)	Serious[1]	N/A	Not serious	Serious[2]	47	MD 0.03 (-0.51, 0.57)	Low
Self-esteem (DQOL) – higher numbers favour intervention							
1 (Cooke)	Serious[1]	N/A	Not serious	Serious[2]	47	MD 0.06 (-0.40, 0.52)	Low
Positive affect (DQOL) – higher numbers favour intervention							
1 (Cooke)	Serious[1]	N/A	Not serious	Serious[2]	47	MD 0.12 (-0.33, 0.57)	Low
Absence of negative affect (DQOL) – higher numbers favour intervention							
1 (Cooke)	Serious[1]	N/A	Not serious	Serious[2]	47	MD 0.04 (-0.33, 0.41)	Low
Feelings of belonging (DQOL) – higher numbers favour intervention							
1 (Cooke)	Serious[1]	N/A	Not serious	Serious[2]	47	MD 0.11 (-0.27, 0.49)	Low
Sense of aesthetics (DQOL) – higher numbers favour intervention							
1 (Cooke)	Serious[1]	N/A	Not serious	Serious[2]	47	MD -0.05 (-0.47, 0.37)	Low
Depression (Geriatric Depression Scale) – lower numbers favour intervention							
1 (Cooke)	Serious[1]	N/A	Not serious	Serious[2]	47	MD 0.24 (-1.46, 1.94)	Low

1. Crossover design with potentially serious confounding.
2. Non-significant result.

Preferred music listening vs usual care

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
Anxiety (RAID) – lower numbers favour intervention							

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
1 (Sung)	Very serious[1]	N/A	Not serious	Serious[2]	52	MD -0.42 (-2.92, 2.08)	Very low

1. Lack of appropriate blinding. Cluster randomised study with only 1 cluster.
2. Non-significant result.

High-intensity exercise vs non-exercise activity program

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
Geriatric Depression Scale (4 months) – lower numbers favour intervention							
1 (Boström)	Not serious	N/A	Not serious	Serious[1]	183	MD -0.05 (-0.84, 0.75)	Moderate
Geriatric Depression Scale (7 months) – lower numbers favour intervention							
1 (Boström)	Not serious	N/A	Not serious	Serious[1]	184	MD -0.06 (-0.89, 0.76)	Moderate
Montgomery-Asberg Depression Rating Scale (4 months) – lower numbers favour intervention							
1 (Boström)	Not serious	N/A	Not serious	Serious[1]	183	MD 0.06 (-1.60, 1.73)	Moderate
Montgomery-Asberg Depression Rating Scale (7 months) – lower numbers favour intervention							
1 (Boström)	Not serious	N/A	Not serious	Serious[1]	184	MD 0.16 (-1.57, 1.89)	Moderate

1. Non-significant result.

G.10.1.2 Antidepressants for other non-cognitive symptoms

SSRIs vs placebo

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
Cohen-Mansfield Agitation Inventory – lower scores favour SSRIs							
3 (Seitz systematic review, Porsteinsson 2014)	Serious[1]	Serious[2]	Not serious	Not serious	419	MD -1.27 (-2.50, -0.03)	Low
NPI – lower scores favour SSRIs							
2 (Finkel 2004, Porsteinsson 2014)	Serious[1]	Serious[2]	Not serious	Serious[3]	409	MD -1.99 (-9.66, 5.68)	Very low
BEHAVE-AD – lower scores favour SSRIs							

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
1 (Finkel 2004)	Serious[1]	N/A	Not serious	Serious[3]	240	MD -0.70 (-1.95, 0.55)	Low
Neurobehavioral Rating Scale – lower scores favour SSRIs							
2 (Pollock 2002, Porsteinsson 2014)	Serious[1]	Serious[2]	Not serious	Serious[3]	219	MD -2.82 (-8.76, 3.13)	Very low
Withdrawal due to adverse events – lower scores favour SSRIs							
4 (Seitz systematic review)	Serious[1]	Not serious	Not serious	Very serious[4]	399	RR 1.15 (0.67, 1.99)	Very low

1. Lack of information on allocation concealment and blinding.
2. i^2 value > 40%.
3. Non-significant result.
4. 95% CI crosses two lines of a defined MID interval

SSRIs vs atypical antipsychotics

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
Neurobehavioral Rating Scale – lower scores favour SSRIs							
1 (Pollock 2007)	Not serious	N/A	Not serious	Serious[1]	103	MD -0.53 (-2.37, 1.31)	Moderate
Neurobehavioral Rating Scale (psychosis subscale) – lower scores favour SSRIs							
1 (Pollock 2007)	Not serious	N/A	Not serious	Serious[1]	103	MD 0.26 (-1.51, 2.03)	Moderate
Withdrawal due to adverse events – lower scores favour SSRIs							
1 (Pollock 2007)	Not serious	N/A	Not serious	Very serious[2]	103	RR 0.42 (0.14, 1.28)	Low

1. Non-significant result.
2. 95% CI crosses two lines of a defined MID interval

SSRIs vs typical antipsychotics

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
Cohen-Mansfield Agitation Inventory – lower scores favour SSRIs							
2 (Seitz systematic review)	Serious[1]	Not serious	Not serious	Serious[2]	33	MD 4.66 (-3.58, 12.90)	Low

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
Neurobehavioral Rating Scale – lower scores favour SSRIs							
1 (Pollock 2002)	Serious[1]	N/A	Not serious	Serious[2]	64	MD -2.80 (-10.34, 4.74)	Low
Withdrawal due to adverse events – lower scores favour SSRIs							
1 (Auchus 1997)	Serious[1]	N/A	Not serious	Very serious[3]	10	RR 0.20 (0.01, 3.35)	Very low

1. Lack of information on allocation concealment and blinding.
2. Non-significant result.
3. 95% CI crosses two lines of a defined MID interval

Trazodone vs placebo

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
Cohen-Mansfield Agitation Inventory – lower scores favour trazodone							
1 (Teri 2000)	Serious[1]	N/A	Not serious	Serious[2]	73	MD 5.18 (-2.86, 13.22)	Low

1. Lack of information on allocation concealment and blinding.
2. Non-significant result.

Trazodone vs typical antipsychotics

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
Cohen-Mansfield Agitation Inventory – lower scores favour trazodone							
2 (Seitz systematic review)	Serious[1]	Not serious	Not serious	Serious[2]	99	MD 3.28 (-3.28, 9.85)	Low

1. Lack of information on allocation concealment and blinding.
2. Non-significant result.

G.10.1.3 Antipsychotics

Atypical antipsychotics vs placebo

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
NPI – lower numbers favours antipsychotics							

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
14 (Ma systematic review)*	Not serious	Not serious	Not serious	Not serious	2,970	MD -2.91 (-4.55, -1.28)	High
Brief psychiatric rating scale – lower numbers favours antipsychotics							
10 (Ma systematic review)*	Not serious	Not serious	Not serious	Not serious	1,957	MD -1.71 (-2.74, -0.68)	High
Cohen-Mansfield Agitation Inventory – lower numbers favours antipsychotics							
8 (Ma systematic review)*	Not serious	Serious[1]	Not serious	Not serious	2,161	MD -1.85 (-3.18, -0.51)	Moderate
Clinical Global Impression of Change – lower numbers favours antipsychotics							
11 (Ma systematic review)*	Not serious	Not serious	Not serious	Not serious	2,566	MD -0.30 (-0.43, -0.18)	High
Adverse events (extrapyramidal) – lower numbers favours antipsychotics							
15 (Ma systematic review)*	Not serious	Not serious	Not serious	Serious[2]	4,092	RR 1.50 (1.24, 1.82)	Moderate
Adverse events (somnolence) – lower numbers favours antipsychotics							
12 (Ma systematic review)*	Not serious	Not serious	Not serious	Not serious	3,838	RR 2.48 (2.00, 3.07)	High
Adverse events (cerebrovascular) – lower numbers favours antipsychotics							
12 (Ma systematic review)*	Not serious	Not serious	Not serious	Serious[2]	3,198	RR 2.24 (1.21, 4.16)	Moderate
Mortality – lower numbers favours antipsychotics							
17 (Ma systematic review)*	Not serious	Not serious	Not serious	Not serious	5,028	RR 1.53 (1.06, 2.22)	High

*Results from the Ma systematic review were converted from odds ratios to relative risks for consistency with the rest of the guideline, and corrections were made where analyses had not correctly accounted for trials with more than 2 arms.

1. $i^2 > 40\%$.
2. 95% CI crosses one line of a defined MID interval

Olanzapine vs haloperidol

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
MMSE – higher numbers favour olanzapine							
1 (Verhey 2006)	Serious[1]	N/A	Not serious	Serious[2]	46	MD 0.66 (-3.79, 5.11)	Low
NPI – lower numbers favour olanzapine							
1 (Verhey 2006)	Serious[1]	N/A	Not serious	Serious[2]	45	MD 7.78 (-5.87, 21.43)	Low
CMAI – lower numbers favour olanzapine							
1 (Verhey 2006)	Serious[1]	N/A	Not serious	Serious[2]	58	MD 6.50 (-2.45, 15.45)	Low

1. Aspects of study design poorly reported.
2. Non-significant result.

Risperidone vs rivastigmine

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
CMAI – lower numbers favour risperidone							
1 (Holmes 2007)	Serious[1]	N/A	Not serious	Not serious	27	MD -22.90 (-36.85, -8.95)	Moderate

1. Aspects of study design poorly reported.

Antipsychotic withdrawal

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
BPSD – lower numbers favour discontinuation							
3 (Pan systematic review)	Not serious	Serious[1]	Not serious	Serious[2]	214	MD 0.19 (-0.20, 0.58)	Low
BPSD worsening – lower numbers favour discontinuation							
7 (Pan systematic review)	Not serious	Not serious	Not serious	Not serious	366	RR 1.78 (1.30, 2.42)	High
Early study termination – lower numbers favour discontinuation							
6 (Pan systematic review)	Not serious	Not serious	Not serious	Serious[3]	462	RR 1.13 (0.88, 1.46)	Moderate
Mortality – lower numbers favour discontinuation							

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
5 (Pan systematic review)	Not serious	Not serious	Not serious	Serious[2]	407	RR 0.79 (0.41, 1.54)	Moderate

1. i^2 value > 40%.
2. Non-significant result.
3. 95% CI crosses one line of a defined MID interval.

Antipsychotic withdrawal UK (6 months)

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
Cognition (SIB) – higher numbers favour continuation							
1 (Ballard 2008)	Not serious	N/A	Not serious	Serious[1]	102	MD -0.4 (-6.4, 5.5)	Moderate
Neuropsychiatric symptoms (NPI) – lower numbers favour continuation							
1 (Ballard 2008)	Not serious	N/A	Not serious	Serious[1]	109	MD -2.4 (-8.2, 3.5)	Moderate
Cognition (MMSE) – higher numbers favour continuation							
1 (Ballard 2008)	Not serious	N/A	Not serious	Serious[1]	84	MD -1.0 (-2.7, 0.7)	Moderate
Parkinsonism (modified UPDRS) – lower numbers favour continuation							
1 (Ballard 2008)	Not serious	N/A	Not serious	Serious[1]	84	MD 1.1 (-0.4, 2.6)	Moderate
Activities of daily living (Bristol ADL) – higher numbers favour continuation							
1 (Ballard 2008)	Not serious	N/A	Not serious	Serious[1]	106	MD 1.7 (-1.2, 4.6)	Moderate
Receptive language (STALD) – higher numbers favour continuation							
1 (Ballard 2008)	Not serious	N/A	Not serious	Serious[1]	73	MD -0.2 (-1.1, 0.6)	Moderate
Expressive skill (STALD) – higher numbers favour continuation							
1 (Ballard 2008)	Not serious	N/A	Not serious	Serious[1]	73	MD -1.0 (-2.0, 0.04)	Moderate
Verbal fluency (FAS) – higher numbers favour continuation							
1 (Ballard 2008)	Not serious	N/A	Not serious	Not serious	56	MD -4.5 (-7.3, -1.7)	High

1. Non-significant result.

Antipsychotic withdrawal UK (12 months)

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
Cognition (SIB) – higher numbers favour continuation							
1 (Ballard 2008)	Not serious	N/A	Not serious	Serious[1]	55	MD -8.4 (-18.6, 1.7)	Moderate
Neuropsychiatric symptoms (NPI) – lower numbers favour continuation							
1 (Ballard 2008)	Not serious	N/A	Not serious	Not serious	59	MD -10.9 (-20.1, -1.7)	High

1. Non-significant result.

Antipsychotic withdrawal UK (24-54 months)

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
Mortality (ITT) – lower numbers favour continuation							
1 (Ballard 2008)	Not serious	N/A	Not serious	Not serious	165	HR 0.58 (0.36, 0.92)	High
Mortality (modified ITT*) – lower numbers favour continuation							
1 (Ballard 2008)	Not serious	N/A	Not serious	Not serious	128	HR 0.58 (0.35, 0.95)	High

*Population restricted to only those individuals who took one dose of study medication

Antipsychotic switch to memantine

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
Bristol Activities of Daily Living score – higher numbers favour memantine							
1 (Ballard 2015)	Not serious	N/A	Not serious	Serious[1]	164	MD 0.23 (-1.80, 2.27)	Moderate
Cohen-Mansfield Agitation Inventory – lower numbers favour memantine							
1 (Ballard 2015)	Not serious	N/A	Not serious	Serious[1]	164	MD 4.09 (-0.35, 8.53)	Moderate
NPI – lower numbers favour memantine							
1 (Ballard 2015)	Not serious	N/A	Not serious	Serious[1]	163	MD 3.63 (-1.40, 8.67)	Moderate
MMSE – higher numbers favour memantine							
1 (Ballard 2015)	Not serious	N/A	Not serious	Serious[1]	113	MD 1.29 (-0.21, 2.79)	Moderate
Serious adverse events – lower numbers favour memantine							
1 (Ballard 2015)	Not serious	N/A	Not serious	Serious[2]	164	RR 0.74 (0.44, 1.24)	Moderate

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
Mortality – lower numbers favour memantine							
1 (Ballard 2015)	Not serious	N/A	Not serious	Serious[1]	164	RR 0.46 (0.15, 1.42)	Moderate

1. Non-significant result
2. 95% CI crosses one line of a defined MID interval
3. 95% CI crosses two lines of a defined MID interval

Enhanced psychosocial care versus usual care

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
Proportion taking neuroleptics – lower numbers favour intervention							
1 (Fossey)	Serious[1]	N/A	Not serious	Not serious	338	RR 0.55 (0.39, 0.76)	Moderate
Fall in past 12 months – lower numbers favour intervention							
1 (Fossey)	Serious[1]	N/A	Not serious	Very serious[3]	340	RR 0.90 (0.59, 1.38)	Very low
Aggression (Cohen-Mansfield agitation score) – lower numbers favour intervention							
1 (Fossey)	Serious[1]	N/A	Not serious	Serious[2]	334	MD 0.3 (-8.3, 8.9)	Low
Wellbeing (dementia care mapping) – higher numbers favour intervention							
1 (Fossey)	Serious[1]	N/A	Not serious	Serious[2]	302	MD -0.2 (-0.5, 0.2)	Low

1. Lack of appropriate blinding
2. Non-significant result.
3. 95% CI crosses two lines of a defined MID interval

G.10.1.4 Memantine vs placebo (mild Alzheimer's disease)

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
ADAS-cog – lower numbers favour intervention							
3 (Schneider systematic review)	Serious[1]	Not serious	Not serious	Serious[2]	425	MD -0.17 (-1.60, 1.26)	Low
ADCS-ADL – lower numbers favour intervention							
3 (Schneider systematic review)	Serious[1]	Not serious	Not serious	Serious[2]	427	MD 0.62 (-1.46, 2.71)	Low

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
NPI – lower numbers favour intervention							
3 (Schneider systematic review)	Serious[1]	Not serious	Not serious	Serious[2]	427	MD 0.09 (-2.11, 2.29)	Low

1. Post-hoc subgroup analysis.
2. Non-significant result.

G.10.1.5 Sleep problems

Melatonin vs placebo

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
Total night-time sleep time (minutes)							
3 (Dowling, Singer, Wade)	Serious[1]	Not serious	Not serious	Serious[4]	195	MD 12.59 (-12.56, 37.74)	Low
Ratio of daytime sleep to night-time sleep							
2 (Dowling, Singer)	Serious[2]	Not serious	Not serious	Serious[4]	184	MD -0.13 (-0.29, 0.03)	Low
Sleep efficiency							
1 (Singer)	Not serious	N/A	Not serious	Serious[4]	151	MD -0.01 (-0.04,0.03)	Moderate
Nocturnal time awake (minutes)							
1 (Singer)	Not serious	N/A	Not serious	Serious[4]	151	MD 9.08 (-7.51, 25.66)	Moderate
Number of night-time awakenings							
1 (Singer)	Not serious	N/A	Not serious	Serious[4]	151	MD 6.00 (-2.65, 14.65)	Moderate
Carer-rated sleep quality, change from baseline							
1 (Singer)	Not serious	N/A	Not serious	Serious[4]	151	MD -0.01 (-0.21, 0.19)	Moderate
Activities of daily living							
1 (Singer)	Not serious	N/A	Not serious	Serious[4]	151	MD 0.40 (-1.41, 2.22)	Moderate
Number of adverse events reported per person							
1 (Singer)	Not serious	N/A	Not serious	Serious[4]	151	MD 0.20 (-0.72, 1.12)	Moderate
Pittsburgh Sleep Quality Index global score							

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
1 (Wade)	Serious[1]	N/A	Serious[3]	Serious[4]	11	MD -1.71 (-4.27,0.87)	Very Low
Pittsburgh Sleep Quality Index sleep latency (minutes)							
1 (Wade)	Serious[1]	N/A	Serious[3]	Serious[4]	11	MD 0.60 (-30.30, 31.50)	Very Low

1. Very high risk of reporting bias for Wade study.
2. Potential problems with sequence generation, allocation concealment and attrition bias.
3. Mean MMSE baseline scores > 20 cut off – patients had mild dementia.
4. Non-significant result

Trazadone vs placebo

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
Total night-time sleep time (minutes)							
1 (Camargos)	Not serious	N/A	Not serious	Not serious	30	MD 42.46 (0.9, 84.0)	High
Sleep efficiency							
1 (Camargos)	Not serious	N/A	Not serious	Not serious	30	MD 8.53 (1.9, 15.1)	High
Nigh-time waking after sleep onset (minutes)							
1 (Camargos)	Not serious	N/A	Not serious	Serious[1]	30	MD -20.41 (-60.4, 19.6)	Moderate
Number of nocturnal awakenings							
1 (Camargos)	Not serious	N/A	Not serious	Serious[1]	30	MD -3.71 (-8.2, 0.8)	Moderate
Total daytime sleep time (minutes)							
1 (Camargos)	Not serious	N/A	Not serious	Serious[1]	30	MD 5.12 (-28.2, 38.4)	Moderate
Number of daytime naps							
1 (Camargos)	Not serious	N/A	Not serious	Serious[1]	30	MD 0.84 (-2.6, 4.3)	Moderate
Activities of daily living (Katz Index)							
1 (Camargos)	Not serious	N/A	Not serious	Serious[1]	30	MD 0.5 (-0.8, 1.8)	Moderate

1. Non-significant result.

Memantine vs placebo

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
Epworth Sleepiness Scale (Scale goes from 0 to 24, higher scores worse)							
1 (Larsson)	Serious[1]	N/A	Not serious	Serious[2]	60	MD -0.35 (-3.26, 2.56)	Low
Stavanger Sleep Questionnaire							
1 (Larsson)	Serious[1]	N/A	Not serious	Not serious	55	MD 0.48 (0.06, 0.90)	Moderate

1. Unclear whether study personnel, medical staff and patients were blinded to treatment and whether placebo and intervention groups were treated equally apart from the intervention.
2. Non-significant result

Light therapy

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
Total sleep duration (minutes, 6-50 days)							
1 (Dowling)	Serious[1]	N/A	Not serious	Serious[3]	35	MD 9.00 (-67.14, 85.14)	Low
Number of night-time awakenings at endpoint							
1 (Dowling)	Serious[1]	N/A	Not serious	Serious[3]	35	MD -4.00 (-11.06, 3.06)	Low
Sleep latency at endpoint (after 3 weeks of treatment)							
1 (Gasio)	Serious[2]	N/A	Not serious	Serious[3]	13	MD -79.00 (-327.17, 169.17)	Low
Sleep latency at follow-up (3 weeks after treatment)							
1 (Gasio)	Serious[2]	N/A	Not serious	Serious[3]	13	MD -62.00 (-216.55, 92.55)	Low
Total sleep duration (minutes) at endpoint (after 3 weeks of treatment)							
1 (Gasio)	Serious[2]	N/A	Not serious	Serious[3]	13	MD 143.00 (-637.66, 923.66)	Low
Total sleep duration (minutes) at follow-up (3 weeks after treatment)							
1 (Gasio)	Serious[2]	N/A	Not serious	Serious[3]	13	MD 110 (-77.22, 297.22)	Low
Night-time activity counts (per night) at endpoint (after 3 weeks of treatment)							
1 (Gasio)	Serious[2]	N/A	Not serious	Serious[3]	13	MD -20.60 (-46.52, 5.32)	Low
Night-time activity counts (per night) at follow-up (3 weeks after treatment)							

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
1 (Gasio)	Serious[2]	N/A	Not serious	Serious[3]	13	MD –24.70 (-52.70, 3.30)	Low

1. Potential problems with sequence generation, allocation concealment and attrition bias.
2. Potential problems with allocation concealment and blinding of assessors.
3. Non-significant result.

Slow-stroke back massage

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
Total night-time sleep time (NTST)							
1 (Harris)	Not serious	N/A	Not serious	Serious[1]	40	MD 35.78 (-12.04, 83.60)	Moderate
Sleep efficiency							
1 (Harris)	Not serious	N/A	Not serious	Serious[1]	40	MD 4.10 (-4.58, 12.78)	Moderate

1. Non-significant result.

Multicomponent non-pharmacological interventions vs usual care

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
Total night-time sleep time (minutes)							
2 (Alessi 2005, McCurry 2011)	Not serious	Not serious	Not serious	Not serious	184	MD 23.72 (0.73, 46.70)	High
Total night-time awake time (minutes)							
2 (McCurry 2005, McCurry 2011)	Not serious	Not serious	Not serious	Not serious	89	MD -38.89 (-65.49, -12.29)	High
Number of night-time awakenings							
3 (Alessi 2005, McCurry 2005, McCurry 2011)	Not serious	Not serious	Not serious	Serious[1]	207	MD -2.20 (-4.83, 0.43)	Moderate
Total daytime sleep time (minutes)							
1 (McCurry 2011)	Not serious	N/A	Not serious	Serious[1]	66	MD -7.30 (-46.82, 32.22)	Moderate
Sleep disorders inventory							
1 (McCurry 2011)	Not serious	N/A	Not serious	Not serious	66	MD -0.90 (-1.45, -0.35)	High

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
RMBPC - depression							
1 (McCurry 2005)	Not serious	N/A	Not serious	Serious[1]	23	MD -0.22 (-0.48, 0.04)	Moderate

1. Non-significant result.
2. Subgroup analyses carried out post-hoc.

Individualised activities

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
Daytime minutes slept							
1 (Richards 2005)	Serious[1]	N/A	Not serious	Not serious	50	MD -45.12 (-72.45, -17.79)	Moderate
Night-time minutes to sleep onset							
1 (Richards 2005)	Serious[1]	N/A	Not serious	Serious[2]	50	MD 9.87 (-18.28, 38.02)	Low
Night-time minutes slept							
1 (Richards 2005)	Serious[1]	N/A	Not serious	Serious[2]	50	MD -4.67 (-74.6, 65.26)	Low
Night-time minutes awake							
1 (Richards 2005)	Serious[1]	N/A	Not serious	Serious[2]	50	MD -21.85 (-94.28, 50.58)	Low
Night-time sleep efficiency							
1 (Richards 2005)	Serious[1]	N/A	Not serious	Serious[2]	50	MD -0.35 (-10.35, 9.65)	Low
Day/night sleep ratio							
1 (Richards 2005)	Serious[1]	N/A	Not serious	Serious[2]	50	MD -0.17 (-0.73, 0.39)	Low

1. Subgroup analyses carried out post-hoc.
2. Non-significant result.

Continuous positive air pressure

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
Epworth Sleepiness Scale 3 weeks (Scale goes from 0 to 24, higher scores worse)							
1 (Chong 2006)	Not Serious	N/A	Not serious	Serious[1]	39	MD -1.10 (-3.10, 0.90)	Moderate

1. Non-significant result.

Non-pharmacological management of agitation, aggression and apathy

Sensory interventions

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
Agitation (CMAI) – lower numbers favour intervention							
5 (Ballard 2002, Yang 2015, Ridder 2013, Lin 2011, Burns 2009)	Not serious	Not serious	Not serious	Serious[1]	446	MD -0.83 (-2.52, 0.85)	Moderate
Negative affect – lower numbers favour intervention							
1 (O'Connor 2013)	Not serious	N/A	Not serious	Serious[1]	64	MD -0.20 (-2.11, 1.71)	Moderate
Positive affect – higher numbers favour intervention							
1 (O'Connor 2013)	Not serious	N/A	Not serious	Serious[1]	64	MD 0.40 (-4.49, 5.29)	Moderate
Agitated behaviours – lower numbers favour intervention							
3 (O'Connor 2013, Sung 2006, Burns 2009)	Not serious	Not serious	Not serious	Serious[2]	141	SMD -0.26 (-0.59, 0.08)	Moderate
Quality of life (ADRQL) - higher numbers favour intervention							
1 (Ridder 2013)	Not serious	N/A	Not serious	Serious[1]	42	MD 17.60 (-24.66, 59.86)	Moderate
Depression (Cornell scale) – lower numbers favour intervention							
1 (Burns 2011)	Not serious	N/A	Not serious	Serious[1]	45	MD 0.50 (-1.15, 2.15)	Moderate
Behavioural pathology (MOUSEPAD, BEHAVE-AD) – lower numbers favour intervention							
2 (Burns 2011, Lyketsos 1999)	Not serious	Not serious	Not serious	Serious[1]	74	MD 0.18 (-0.27, 0.64)	Moderate
MMSE – higher numbers favour intervention							
1 (Burns 2011)	Not serious	N/A	Not serious	Serious[1]	46	MD 1.80 (-1.41, 5.01)	Moderate

1. Non-significant result.
2. 95% CI crosses one line of a defined MID interval.

Social contact

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
Agitation – lower numbers favour intervention							
2 (Camberg 1999, Churchill 1999)	Not serious	Serious[1]	Not serious	Very serious[2]	164	SMD -0.19 (-0.71, 0.33)	Very low

1. $i^2 > 40\%$.
2. 95% CI crosses two lines of a defined MID interval.

Activities

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
Agitation – lower numbers favour intervention							
6 (C-M 2007, C-M 2012, Fitzsimmons 2002, Kolanowski 2001, van der Ploeg 2013, Watson 1998)	Serious[3]	Serious[1]	Not serious	Serious[4]	465	SMD -0.34 (-0.74, 0.05)	Very low
Negative affect – lower numbers favour intervention							
3 (C-M 2007, C-M 2012, van der Ploeg 2013)	Serious[3]	Not serious	Not serious	Not serious	336	MD -0.02 (-0.04, -0.00)	Moderate
Pleasurable affect – higher numbers favour intervention							
3 (C-M 2007, C-M 2012)	Serious[3]	Serious[1]	Not serious	Not serious	292	MD 0.29 (0.15, 0.42)	Low
Interested affect – higher numbers favour intervention							
3 (C-M 2007, C-M 2012, van der Ploeg 2013)	Serious[3]	Serious[1]	Not serious	Not serious	336	SMD 0.57 (0.23, 0.90)	Low
Constructive engagement – higher numbers favour intervention							
1 (van der Ploeg 2013)	Serious[3]	N/A	Not serious	Serious[2]	44	MD 0.30 (-2.32, 2.92)	Low
Negative engagement – lower numbers favour intervention							

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
1 (van der Ploeg 2013)	Serious[3]	N/A	Not serious	Serious[2]	44	MD -0.20 (-5.46, 5.06)	Low

1. i^2 > 40%.
2. Non-significant result.
3. Methods of randomisation unclear
4. 95% CI crosses one line of a defined MID interval.

Care delivery interventions

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
Agitation (CMAI) – lower numbers favour intervention							
2 (Rapp 2013, Zwijsen 2014)	Not serious	Serious[1]	Not serious	Serious[2]	701	MD -6.06 (-14.04, 1.92)	Low
Aggressive behaviours – lower numbers favour intervention							
2 (Rapp 2013, Zwijsen 2014)	Not serious	Serious[1]	Not serious	Very serious[3]	701	SMD -0.30 (-0.99, 0.38)	Very low
Number of psychotropic prescriptions							
1 (Rapp 2013)	Not serious	N/A	Not serious	Serious[2]	304	MD -0.03 (-0.13, 0.07)	Moderate
Number of antidepressant prescriptions							
1 (Rapp 2013)	Not serious	N/A	Not serious	Not serious	304	MD 0.04 (0.03, 0.05)	Moderate
Number of cholinesterase inhibitor prescriptions							
1 (Rapp 2013)	Not serious	N/A	Not serious	Not serious	304	MD 0.11 (0.10, 0.12)	Moderate

1. i^2 > 40%.
2. Non-significant result.
3. 95% CI crosses two lines of a defined MID interval.

Staff training

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
Agitation (CMAI) – lower numbers favour intervention							
1 (Deudon 2009)	Not serious	N/A	Not serious	Not serious	272	MD -5.69 (-9.85, -1.53)	High

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
Physically aggressive behaviours – lower numbers favour intervention							
1 (Deudon 2009)	Not serious	N/A	Not serious	Serious[1]	272	MD -0.08 (-0.39, 0.23)	Moderate
Verbally aggressive behaviours – lower numbers favour intervention							
1 (Deudon 2009)	Not serious	N/A	Not serious	Not serious	272	MD -0.16 (-0.32, -0.00)	High

1. Non-significant result.

Gingko biloba

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
NPI total score – lower numbers favour intervention							
4 (Herrschaft 2012, Ihl 2011, Napryeyenko 2007, Nikolova 2013)	Not serious	Serious[1]	Not serious	Not serious	1,596	MD -3.86 (-7.62, -0.10)	Moderate
NPI distress score – lower numbers favour intervention							
4 (Herrschaft 2012, Ihl 2011, Napryeyenko 2007, Nikolova 2013)	Not serious	Serious[1]	Not serious	Not serious	1,596	MD -2.33 (-4.34, -0.33)	Moderate
Activities of daily living – lower numbers favour intervention							
4 (Herrschaft 2012, Ihl 2011, Napryeyenko 2007, Nikolova 2013)	Not serious	Serious[1]	Not serious	Serious[2]	1,596	SMD -0.54 (-0.91, -0.18)	Low
Quality of life – higher numbers favour intervention							
2 (Herrschaft 2012, Ihl 2011)	Not serious	Not serious	Not serious	Not serious	806	MD 2.00 (0.88, 3.12)	High
Clinical global assessment – lower numbers favour intervention							
4 (Herrschaft 2012, Ihl 2011,	Not serious	Serious[1]	Not serious	Not serious	1,590	MD -0.75 (-1.34, -0.15)	Moderate

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
Napryeyenko 2007, Nikolova 2013)							
Cognition – lower numbers favour intervention							
4 (Herrschaft 2012, Ihl 2011, Napryeyenko 2007, Nikolova 2013)	Not serious	Serious[1]	Not serious	Serious[2]	1,590	SMD -0.78 (-1.50, -0.05)	Low

1. $i^2 > 40\%$.
2. 95% CI crosses one line of a defined MID interval

G.10.1.6 Pharmacological management of agitation, aggression and apathy

Mood stabilisers vs placebo

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
Agitation: CMAI – lower numbers favour mood stabilisers							
4 (Herrmann 2007, Porsteinsson 2001, Profenno 2005, Tariot 2005)	Not serious	Serious[1]	Not serious	Serious[2]	254	MD -0.67 (-3.42, 4.77)	Low
NPI/BPRS subscale agitation/aggression - lower numbers favour mood stabilisers							
2 (Herrmann 2007, Tariot 2005)	Not serious	Serious[1]	Not serious	Very serious[3]	172	SMD 0.40 (-0.31, 1.10)	Very low
Neuropsychiatric profile NPI total score - lower numbers favour mood stabilisers							
2 (Herrmann 2007, Profenno 2005)	Not serious	Not serious	Not serious	Not Serious	51	MD 2.87 (1.01, 4.73)	High
Brief Psychiatric Rating scale - lower numbers favour mood stabilisers							
2 (Porsteinsson 2001, Tariot 2005, Olin 2001)	Not serious	Not serious	Not serious	Serious[2]	224	MD 0.46 (-1.78, 2.70)	Moderate
Physical Self Maintenance Scale – lower numbers favour mood stabilisers							

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
4 (Porsteinsson 2001, Profenno 2005, Tariot 2005, Olin 2001)	Not serious	Not serious	Not serious	Serious[2]	248	MD 0.15 (-0.27, 0.57)	Moderate
Cognition MMSE – higher numbers favours mood stabilisers							
4 (Herrmann; Porsteinsson; Tariot; Olin)	Not serious	Not serious	Not serious	Not serious	273	MD -0.94 (-1.72, -0.17)	High
Any adverse events - lower numbers favour mood stabilisers							
2 (Herrmann 2007, Porsteinsson 2001)	Not serious	Not serious	Not serious	Serious[4]	83	RR 1.77 (1.19, 2.62)	Moderate
Serious adverse events - lower numbers favour mood stabilisers							
1 (Porsteinsson 2001)	Not serious	N/A	Not serious	Very serious[3]	56	RR 1.00 (0.15, 6.61)	Low

1. i^2 value > 40%.
2. Non-significant result.
3. 95% CI crosses two lines of a defined MID interval
4. 95% CI crosses one line of a defined MID interval

Cholinesterase inhibitors vs placebo

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
Agitation: CMAI – lower numbers favour cholinesterase inhibitors							
1 (Howard 2007)	Not serious	N/A	Not serious	Serious[1]	221	MD 1.35 (-3.85, 6.54)	Moderate
Neuropsychiatric profile NPI total score - lower numbers favour cholinesterase inhibitors							
3 (Holmes 2004, Howard 2007, Mahlberg 2007)	Not serious	Serious[2]	Not serious	Serious[1]	317	MD -4.95 (-11.19, 1.29)	Low
Neuropsychiatric profile NPI agitation subscale – lower numbers favour cholinesterase inhibitors							
1 (Mahlberg 2007)	Not serious	N/A	Not serious	Not serious	20	MD -5.20 (-7.95, -2.45)	Moderate
Global assessment SIB - higher numbers favour cholinesterase inhibitors							

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
1 (Howard 2007)	Not serious	N/A	Not serious	Not serious	60	MD 6.75 (1.59, 11.91)	High
NOSGER- higher favours cholinesterase inhibitors							
1 (Mahlberg 2007)	Not serious	N/A	Not serious	Serious[1]	20	MD -6.60 (-23.30, 10.10)	Moderate
Cognition (standardised MMSE) higher favours cholinesterase inhibitors							
1 (Howard 2007)	Not serious	N/A	Not serious	Not serious	113	MD 1.50 (0.15, 2.85)	High

1. Non-significant result.
2. i^2 value > 40%.

Memantine vs placebo

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
Agitation: CMAI – lower numbers favour memantine							
1 (Fox 2012)	Not serious	N/A	Not serious	Serious[1]	149	MD -3.10 (-9.43, 3.23)	Moderate
Neuropsychiatric profile NPI total score - lower numbers favour memantine							
1 (Fox 2012)	Not serious	N/A	Not serious	Not serious	138	MD -9.40 (-15.41, -3.39)	High
Global assessment SIB - higher numbers favour memantine							
1 (Fox 2012)	Not serious	N/A	Not serious	Serious[1]	149	MD 2.40 (-1.81, 6.61)	Moderate
Clinicians global impression of change CGIC - higher numbers favour memantine							
1 (Fox 2012)	Not serious	N/A	Not serious	Serious[1]	149	MD -0.10 (-0.60, 0.40)	Moderate
Cognition (standardised MMSE) – higher numbers favour memantine							
1 (Fox 2012)	Not serious	N/A	Not serious	Serious[1]	149	MD 1.00 (-1.16, 3.16)	Moderate

1. Non-significant result.

Tetrahydrocannabinol vs placebo

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
Agitation CMAI – lower numbers favour THC							
1 (van den Elsen 2015)	Not serious	N/A	Not serious	Serious[1]	47	MD 2.80 (-7.43, 13.03)	Moderate
Neuropsychiatric profile NPI total score - lower numbers favour THC							

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
1 (van den Elsen 2015)	Not serious	N/A	Not serious	Serious[1]	47	MD 3.90 (-4.69, 12.49)	Moderate
NPI agitation/aggression subscale – lower numbers favour THC							
1 (van den Elsen 2015)	Not serious	N/A	Not serious	Serious[1]	47	MD 0.10 (-2.30, 2.50)	Moderate
NPI aberrant behaviour subscale – lower numbers favour THC							
1 (van den Elsen 2015)	Not serious	N/A	Not serious	Serious[1]	47	MD -0.10 (-2.45, 2.25)	Moderate
Caregivers Clinical global impression of change CCGIC- higher numbers favour THC							
1 (van den Elsen 2015)	Not serious	N/A	Not serious	Serious[1]	46	MD 0.30 (-0.48, 1.08)	Moderate
Activities of daily living - Barthel index- higher numbers favour THC							
1 (van den Elsen 2015)	Not serious	N/A	Not serious	Serious[1]	46	MD 1.30 (-1.73, 4.33)	Moderate
Quality of life QoL AD – higher numbers favour THC							
1 (van den Elsen 2015)	Not serious	N/A	Not serious	Serious[1]	43	MD -1.60 (-4.47, 1.27)	Moderate

1. Non-significant result.

Prazosin vs placebo

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
Neuropsychiatric profile NPI total score - lower numbers favour prazosin							
1 (Wang 2008)	Very serious[1]	N/A	Not serious	Serious[2]	13	MD -18.00 (-41.93, 5.93)	Very low
Brief Psychiatric rating scale – lower numbers favour prazosin							
1 (Wang 2008)	Very serious[1]	N/A	Not serious	Not serious	13	MD -12.00 (-19.15, -4.85)	Low
Clinicians global impression of change CGIC - higher numbers favour prazosin							
1 (Wang 2008)	Very serious[1]	N/A	Not serious	Not serious	13	MD -1.90 (-3.38, -0.42)	Low

1. Study at high risk of bias.
2. Non-significant result.

Dextromethorphan-quinidine vs placebo

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
NPI – lower numbers favour intervention							
1 (Cummings 2015)	Not serious	N/A	Not serious	Not serious	159	MD -5.90 (-11.68, -0.12)	High
NPI agitation/aggression subscale – lower numbers favour intervention							
1 (Cummings 2015)	Not serious	N/A	Not serious	Not serious	159	MD -1.70 (-2.84, -0.56)	High
Depression (Cornell scale) – lower numbers favour intervention							
1 (Cummings 2015)	Not serious	N/A	Not serious	Not serious	141	MD -1.60 (-2.92, -0.28)	High
CGIC – higher numbers favour intervention							
1 (Cummings 2015)	Not serious	N/A	Not serious	Serious[1]	152	MD 1.00 (-1.06, 3.06)	Moderate
MMSE – higher numbers favour intervention							
1 (Cummings 2015)	Not serious	N/A	Not serious	Serious[1]	151	MD 0.70 (-0.41, 1.81)	Moderate
QoL ADS – higher numbers favour intervention							
1 (Cummings 2015)	Not serious	N/A	Not serious	Serious[1]	152	MD 0.40 (-1.42, 2.22)	Moderate
Any adverse events – lower numbers favour intervention							
1 (Cummings 2015)	Not serious	N/A	Not serious	Not serious	279	RR 1.41 (1.12, 1.79)	High
Serious adverse events – lower numbers favour intervention							
1 (Cummings 2015)	Not serious	N/A	Not serious	Serious[1]	279	RR 1.67 (0.65, 4.33)	Moderate
Mortality – lower numbers favour intervention							
1 (Cummings 2015)	Not serious	N/A	Not serious	Very serious[2]	279	No deaths in either arm	Low

1. Non-significant result.
2. Relative risk could not be calculated.

Modafinil vs placebo

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
FrsBe Apathy – lower numbers favour modafinil							
1 (Frakey 2012)	Not serious	N/A	Not serious	Serious[1]	22	MD 7.00 (-2.80, 16.80)	Moderate
DAFS functional assessment – higher numbers favour modafinil							

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
1 (Frakey 2012)	Not serious	N/A	Not serious	Serious[1]	22	MD -3.09 (-12.80, 6.62)	Moderate
Activities of daily living – higher numbers favour modafinil							
1 (Frakey 2012)	Not serious	N/A	Not serious	Serious[1]	22	MD -3.36 (-7.74, 1.02)	Moderate
Zarit carer burden index – lower numbers favour modafinil							
1 (Frakey 2012)	Not serious	N/A	Not serious	Serious[1]	22	MD 0.00 (-12.40, 12.40)	Moderate

1. Non-significant result.

Donepezil and choline alphoscerate vs donepezil

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
FrsBe Apathy severity- lower numbers favour donepezil and choline							
1 (Rea 2015)	Not serious	N/A	Not serious	Not serious	113	MD -2.70 (-4.69, -0.71)	High
NPI severity - lower numbers favour donepezil and choline							
1 (Rea 2015)	Not serious	N/A	Not serious	Not serious	113	MD -7.70 (-14.23, -1.17)	High
Frontal Assessment Battery – higher numbers favour donepezil and choline							
1 (Rea 2015)	Not serious	N/A	Not serious	Not serious	113	MD 1.60 (0.48, 2.72)	High
MMSE – higher numbers favour donepezil and choline							
1 (Rea 2015)	Not serious	N/A	Not serious	Not serious	113	MD 2.50 (0.59, 4.41)	High
1 ADAS cog –lower numbers favour donepezil and choline							
1 (Rea 2015)	Not serious	N/A	Not serious	Not serious	113	MD -8.50 (-13.65, -3.35)	High

G.11 Supporting informal carers

G.11.1 Supporting informal carers of people living with dementia

- How effective are carers' assessments in identifying the needs of informal carers of people living with dementia?
- What interventions/services are most effective for supporting the wellbeing of informal carers of people living with dementia?

G.11.1.1 Psychoeducational interventions

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Usual care	Summary of results	
Carer burden (lower values favour intervention)									
3	RCT	Serious[1]	Not serious	Not serious	Serious[2]	201	172	SMD -0.14 (-0.34, 0.07)	Low
Carer depression (lower values favour intervention)									
3	RCT	Not serious	Not serious	Serious[3]	Very serious[4]	192	185	SMD -0.02 (-0.31, 0.28)	Very low
Carer anxiety (lower values favour intervention)									
2	RCT	Not serious	Not serious	Not serious	Serious[2]	151	96	SMD -0.08 (-0.34, 0.18)	Moderate
Carer stress (lower values favour intervention)									
2	RCT	Not serious	Not serious	Not serious	Very serious[4]	41	31	SMD -0.20 (-0.67, 0.28)	Low
Carer quality of life (higher values favour intervention)									
1 (Hattink 2015)	RCT	Not serious	Not serious	N/A	Very serious[4]	21	25	SMD 0.34 (-0.25, 0.92)	Low
Carer self-efficacy (higher values favour intervention)									
3	RCT	Serious[4]	Not serious	Not serious	Serious[2]	174	159	SMD 0.20 (-0.02, 0.41)	Low
Carer social support (higher values favour intervention)									
1 (Hebert 2003)	RCT	Not serious	Not serious	N/A	Very serious[4]	60	56	SMD 0.04 (-0.33, 0.40)	Low
Revised memory and behaviour problems checklist – severity (lower values favour intervention)									
2	RCT	Not serious	Not serious	Serious[3]	Very serious[4]	153	134	SMD -0.04 (-0.75, 0.67)	Very low
Revised memory and behaviour problems checklist – reaction (lower values favour intervention)									
2	RCT	Not serious	Not serious	Not serious	Serious[2]	153	134	SMD -0.16 (-0.40, 0.07)	Moderate
Activities of daily living – person living with dementia (higher values favour intervention)									
1 (Gitlin 2001)	RCT	Not serious	Not serious	N/A	Serious[2]	93	78	SMD 0.22 (-0.08, 0.52)	Moderate

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Usual care	Summary of results	
Proportion entering long stay care (lower values favour intervention)									
1 (Nobili 2004)	RCT	Not serious	Not serious	N/A	Serious[2]	156	136	RR 1.29 (0.80, 2.08)	Moderate

1. Unclear reporting of methods
2. Crosses one line of a defined MID
3. i^2>40%
4. Crosses two lines of a defined MID

G.11.1.2 Skills training

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Usual care	Summary of results	
Carer burden (lower values favour intervention)									
6	RCT	Serious[1]	Not serious	Not serious	Serious[2]	198	162	SMD -0.36 (-0.57, -0.15)	Low
Carer depression (lower values favour intervention)									
8	RCT	Not serious	Not serious	Not serious	Serious[2]	279	217	SMD -0.16 (-0.34, 0.03)	Moderate
Carer anxiety (lower values favour intervention)									
4	RCT	Not serious	Not serious	Serious[3]	Serious[2]	170	159	SMD -0.22 (-0.62, 0.19)	Low
Carer stress (lower values favour intervention)									
2	RCT	Not serious	Not serious	Not serious	Very serious[4]	40	25	SMD -0.16 (-0.67, 0.35)	Low
Carer quality of life (higher values favour intervention)									
1 (Martin-Carrasco 2009)	RCT	Not serious	Not serious	N/A	Serious[2]	44	38	SMD 0.52 (0.08, 0.96)	Moderate
Carer self-efficacy (higher values favour intervention)									
3	RCT	Not serious	Not serious	Not serious	Serious[2]	103	89	SMD 0.23 (-0.05, 0.52)	Moderate
Carer social support (higher values favour intervention)									
1 (Burgio 2003)	RCT	Serious[3]	Not serious	N/A	Very serious[4]	53	53	SMD 0.06 (-0.32, 0.44)	Very low
Revised memory and behaviour problems checklist – severity (lower values favour intervention)									
4	RCT	Not serious	Not serious	Not serious	Serious[2]	189	148	SMD -0.19 (-0.41, 0.03)	Moderate
Revised memory and behaviour problems checklist – reaction (lower values favour intervention)									
3	RCT	Not serious	Not serious	Serious[2]	Very serious[4]	120	91	SMD -0.16 (-0.55, 0.22)	Very low

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Usual care	Summary of results	
Behavioural and psychological symptoms of dementia – severity (lower values favour intervention)									
1 (Oken 2010)	RCT	Not serious	Not serious	N/A	Very serious[4]	11	10	SMD 0.46 (-0.61, 1.33)	Low
Behavioural and psychological symptoms of dementia – reaction (lower values favour intervention)									
1 (Zarit 1982)	RCT	Not serious	Not serious	N/A	Very serious[4]	11	10	SMD -0.42 (-1.29, 0.45)	Low

1. Unclear reporting of methods
2. Crosses one line of a defined MID
3. i^2>40%
4. Crosses two lines of a defined MID

G.11.1.3 Psychoeducation and skills training

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Usual care	Summary of results	
Carer burden (lower values favour intervention)									
10	RCT	Not serious	Not serious	Serious[1]	Serious[2]	595	551	SMD -0.30 (-0.49, -0.10)	Low
Carer depression (lower values favour intervention)									
14	RCT	Not serious	Not serious	Not serious	Serious[2]	1,102	929	SMD -0.25 (-0.33, -0.16)	Moderate
Carer anxiety (lower values favour intervention)									
6	RCT	Not serious	Not serious	Not serious	Serious[2]	606	483	SMD -0.26 (-0.39, -0.14)	Moderate
Carer stress (lower values favour intervention)									
6	RCT	Not serious	Not serious	Not serious	Serious[2]	323	323	SMD -0.21 (-0.37, -0.06)	Moderate
Carer quality of life (higher values favour intervention)									
5	RCT	Not serious	Not serious	Serious[1]	Serious[2]	334	324	SMD 0.11 (-0.11, 0.33)	Low
Carer self-efficacy (higher values favour intervention)									
7	RCT	Not serious	Not serious	Serious[1]	Serious[2]	503	470	SMD 0.20 (-0.01, 0.42)	Low
Revised memory and behaviour problems checklist – severity (lower values favour intervention)									
3	RCT	Not serious	Not serious	Serious[1]	Very serious[3]	115	92	SMD -0.11 (-0.52, 0.30)	Very low
Revised memory and behaviour problems checklist – reaction (lower values favour intervention)									
2	RCT	Not serious	Not serious	Serious[1]	Serious[2]	211	172	SMD -0.24 (-0.54, 0.07)	Low
Behavioural and psychological symptoms of dementia – severity (lower values favour intervention)									

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Usual care	Summary of results	
7	RCT	Not serious	Not serious	Serious[1]	Serious[2]	295	289	SMD -0.27 (-0.53, -0.02)	Low
Behavioural and psychological symptoms of dementia – reaction (lower values favour intervention)									
3	RCT	Not serious	Not serious	Not serious	Serious[2]	68	74	SMD -0.23 (-0.56, 0.10)	Moderate
Activities of daily living – person living with dementia (higher values favour intervention)									
3	RCT	Not serious	Not serious	Not serious	Serious[2]	128	133	SMD -0.07 (-0.31, 0.18)	Moderate
Proportion entering long stay care (lower values favour intervention)									
3	RCT	Not serious	Not serious	Not serious	Serious[2]	265	195	RR 1.47 (0.91, 2.37)	Moderate

1. i^2>40%
2. Crosses one line of a defined MID
3. Crosses two lines of a defined MID

G.11.1.4 Supportive interventions

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Usual care	Summary of results	
Carer burden (lower values favour intervention)									
5	RCT	Not serious	Not serious	Not serious	Serious[1]	166	165	SMD -0.10 (-0.31, 0.12)	Moderate
Carer depression (lower values favour intervention)									
5	RCT	Not serious	Not serious	Serious[2]	Serious[1]	240	235	SMD -0.21 (-0.51, 0.10)	Low
Carer anxiety (lower values favour intervention)									
3	RCT	Not serious	Not serious	Serious[2]	Very serious[3]	61	58	SMD 0.08 (-0.63, 0.79)	Very low
Carer stress (lower values favour intervention)									
1 (Quayhagen 2000)	RCT	Not serious	Not serious	N/A	Very serious[3]	22	15	SMD -0.36 (-1.03, 0.30)	Low
Carer quality of life (higher values favour intervention)									
2	RCT	Not serious	Not serious	Serious[2]	Very serious[3]	121	132	SMD 1.34 (-0.91, 3.60)	Very low
Carer social support (higher values favour intervention)									
2	RCT	Not serious	Not serious	Not serious	Very serious[3]	123	138	SMD -0.02 (-0.26, 0.23)	Low
Revised memory and behaviour problems checklist – severity (lower values favour intervention)									
3	RCT	Not serious	Not serious	Not serious	Very serious[3]	72	70	SMD 0.04 (-0.29, 0.37)	Low

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Usual care	Summary of results	

1. Crosses one line of a defined MID
2. i^2>40%
3. Crosses two lines of a defined MID

G.11.1.5 Respite care

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Usual care	Summary of results	
Carer burden versus usual care (lower values favour intervention)									
1 (Wishart 2000)	RCT	Not serious	Not serious	N/A	Very serious[1]	11	10	SMD -0.67 (-1.55, 0.22)	Low
Carer depression versus usual care (lower values favour intervention)									
1 (Grant 2003)	RCT	Not serious	Not serious	N/A	Very serious[1]	32	23	SMD -0.03 (-0.56, 0.51)	Low
Carer depression versus polarity therapy (lower values favour intervention)									
1 (Korn 2009)	RCT	Not serious	Serious[2]	N/A	Serious[3]	18	20	SMD 0.66 (0.01, 1.32)	Low
Carer anxiety versus usual care (lower values favour intervention)									
1 (Grant 2003)	RCT	Not serious	Not serious	N/A	Very serious[1]	32	23	SMD 0.01 (-0.53, 0.54)	Low
Carer stress versus polarity therapy (lower values favour intervention)									
1 (Korn 2009)	RCT	Not serious	Serious[2]	N/A	Serious[3]	18	20	SMD 0.82 (0.15, 1.48)	Low

1. Crosses two lines of a defined MID
2. Polarity therapy not a relevant comparator for the UK
3. Crosses one line of a defined MID

G.11.1.6 Psychotherapy

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Usual care	Summary of results	
Carer burden (lower values favour intervention)									
2	RCT	Not serious	Not serious	Not serious	Not serious	57	50	SMD -0.82 (-1.22, -0.42)	High
Carer depression (lower values favour intervention)									
14	RCT	Serious[1]	Not serious	Serious[2]	Not serious	491	543	SMD -0.55 (-0.85, -0.26)	Low
Carer anxiety (lower values favour intervention)									

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Usual care	Summary of results	
3	RCT	Serious[1]	Not serious	Not serious	Serious[3]	106	122	SMD -0.43 (-0.70, -0.17)	Low
Carer stress (lower values favour intervention)									
3	RCT	Serious[1]	Not serious	Not serious	Serious[3]	158	140	SMD -0.17 (-0.40, 0.06)	Low
Carer quality of life (higher values favour intervention)									
2	RCT	Not serious	Not serious	Not serious	Serious[3]	85	87	SMD 0.35 (0.05, 0.66)	Moderate
Carer self-efficacy (higher values favour intervention)									
4	RCT	Not serious	Not serious	Serious[2]	Serious[3]	82	87	SMD 1.03 (0.05, 2.01)	Low
Revised memory and behaviour problems checklist – severity (lower values favour intervention)									
2	RCT	Not serious	Not serious	Serious[2]	Very serious[4]	82	91	SMD -0.14 (-0.63, 0.34)	Very low
Revised memory and behaviour problems checklist – reaction (lower values favour intervention)									
3	RCT	Not serious	Not serious	Not serious	Serious[3]	167	161	SMD -0.28 (-0.50, -0.07)	Moderate

1. Unclear reporting of methods
2. i^2>40%
3. Crosses one line of a defined MID
4. Crosses two lines of a defined MID

G.11.1.7 Case management

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Usual care	Summary of results	
Carer burden (lower values favour intervention)									
2	RCT	Not serious	Not serious	Not serious	Very serious[1]	98	70	SMD -0.06 (-0.37, 0.25)	Low
Carer depression (lower values favour intervention)									
2	RCT	Not serious	Not serious	Serious[2]	Very serious[1]	98	70	SMD -0.19 (-0.61, 0.23)	Very low
Carer anxiety (lower values favour intervention)									
1 (Xiao 2016)	RCT	Not serious	Not serious	N/A	Serious[3]	31	30	SMD -0.70 (-1.22, -0.18)	Moderate
Carer quality of life (higher values favour intervention)									
1 (Jansen 2011)	RCT	Not serious	Not serious	N/A	Serious[3]	54	45	SMD 0.23 (-0.17, 0.62)	Moderate
Carer self-efficacy (higher values favour intervention)									
3	RCT	Not serious	Not serious	Serious[2]	Very serious[1]	129	100	SMD 0.34 (-0.64, 1.31)	Very low

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Usual care	Summary of results	
Behavioural and psychological symptoms of dementia – severity (lower values favour intervention)									
1 (Xiao 2016)	RCT	Not serious	Not serious	N/A	Serious[3]	31	30	SMD -0.63 (-1.15, -0.12)	Moderate
Proportion entering long stay care (lower values favour intervention)									
1 (Fortinsky 2009)	RCT	Not serious	Not serious	N/A	Serious[3]	44	25	RR 0.41 (0.14, 1.15)	Moderate

1. Crosses two lines of a defined MID
2. i^2>40%
3. Crosses one line of a defined MID

G.11.1.8 Multicomponent interventions

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Usual care	Summary of results	
Carer burden (lower values favour intervention)									
15	RCT	Not serious	Not serious	Serious[1]	Serious[2]	1,663	1,581	SMD -0.17 (-0.33, -0.01)	Low
Carer depression (lower values favour intervention)									
20	RCT	Not serious	Not serious	Serious[1]	Serious[2]	2,806	2,414	SMD -0.29 (-0.49, -0.09)	Low
Carer anxiety (lower values favour intervention)									
2	RCT	Not serious	Not serious	Not serious	Very serious[3]	43	35	SMD 0.05 (-0.40, 0.50)	Low
Carer quality of life (higher values favour intervention)									
3	RCT	Not serious	Not serious	Serious[1]	Serious[2]	337	343	SMD 0.34 (0.04, 0.64)	Low
Carer self-efficacy (higher values favour intervention)									
1 (Martin-Cook 2005)	RCT	Not serious	Not serious	N/A	Very serious[3]	24	23	SMD 0.24 (-0.34, 0.81)	Low
Carer social support (higher values favour intervention)									
2	RCT	Not serious	Not serious	Not serious	Not serious	60	62	SMD 0.56 (0.20, 0.92)	High
Revised memory and behaviour problems checklist – severity (lower values favour intervention)									
4	RCT	Not serious	Not serious	Not serious	Serious[2]	805	549	SMD -0.12 (-0.23, -0.01)	Moderate
Revised memory and behaviour problems checklist – reaction (lower values favour intervention)									
4	RCT	Not serious	Not serious	Not serious	Serious[2]	282	272	SMD -0.19 (-0.43, 0.06)	Moderate

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Usual care	Summary of results	
Behavioural and psychological symptoms of dementia – severity (lower values favour intervention)									
8	RCT	Not serious	Not serious	Serious[1]	Serious[2]	465	479	SMD -0.29 (-0.64, 0.07)	Low
Behavioural and psychological symptoms of dementia – reaction (lower values favour intervention)									
6	RCT	Not serious	Not serious	Not serious	Serious[2]	391	409	SMD -0.31 (-0.45, -0.18)	Moderate
Activities of daily living – person living with dementia (higher values favour intervention)									
6	RCT	Not serious	Not serious	Serious[1]	Serious[2]	430	455	SMD 0.33 (-0.15, 0.81)	Low
Proportion entering long stay care (lower values favour intervention)									
7	RCT	Not serious	Not serious	Serious[1]	Serious[2]	520	472	RR 0.80 (0.61, 1.04)	Low

1. i^2>40%
2. Crosses one line of a defined MID
3. Crosses two lines of a defined MID

G.11.1.9 Exercise

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Usual care	Summary of results	
Carer burden (lower values favour intervention)									
2	RCT	Not serious	Not serious	Serious[1]	Very serious[2]	86	75	SMD -1.76 (-5.27, 1.75)	Very low
Carer depression (lower values favour intervention)									
2	RCT	Not serious	Not serious	Serious[1]	Very serious[2]	86	75	SMD -0.47 (-2.02, 1.09)	Very low
Carer stress (lower values favour intervention)									
1 (Connell 2009)	RCT	Not serious	Not serious	N/A	Serious[2]	69	61	SMD 0.17 (-0.18, 0.51)	Moderate

1. i^2>40%
2. Crosses two lines of a defined MID
3. Crosses one line of a defined MID

G.11.1.10 Memory clinic

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Usual care	Summary of results	
Carer burden (lower values favour intervention)									

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Usual care	Summary of results	
1 (Logiudice 1999)	RCT	Not serious	Not serious	N/A	Very serious[1]	16	14	SMD -0.30 (-1.03, 0.42)	Low
Revised memory and behaviour problems checklist – reaction (lower values favour intervention)									
1 (Logiudice 1999)	RCT	Not serious	Not serious	N/A	Very serious[1]	15	12	SMD 0.15 (-0.61, 0.91)	Low

1. Crosses two lines of a defined MID

G.11.1.11 Meditation/mindfulness

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Usual care	Summary of results	
Carer burden (lower values favour intervention)									
1 (Whitebird 2012)	RCT	Not serious	Not serious	N/A	Very serious[1]	35	35	SMD -0.10 (-0.56, 0.37)	Low
Carer depression (lower values favour intervention)									
5	RCT	Not serious	Not serious	Not serious	Serious[2]	101	91	SMD -0.48 (-0.77, -0.19)	Moderate
Carer anxiety (lower values favour intervention)									
3	RCT	Not serious	Not serious	Serious[3]	Serious[2]	68	65	SMD -0.72 (-1.57, 0.14)	Low
Carer stress (lower values favour intervention)									
3	RCT	Not serious	Not serious	Not serious	Serious[2]	53	54	SMD -0.22 (-0.60, 0.17)	Moderate
Carer self-efficacy (higher values favour intervention)									
1 (Oken 2010)	RCT	Not serious	Not serious	N/A	Very serious[1]	10	10	SMD 0.00 (-0.88, 0.88)	Low
Carer social support (higher values favour intervention)									
1 (Whitebird 2012)	RCT	Not serious	Not serious	N/A	Very serious[1]	35	35	SMD 0.06 (-0.41, 0.52)	Low
Revised memory and behaviour problems checklist – reaction (lower values favour intervention)									
1 (Oken 2010)	RCT	Not serious	Not serious	N/A	Very serious[1]	10	10	SMD -0.08 (-0.96, 0.80)	Low
Behavioural and psychological symptoms of dementia – severity (lower values favour intervention)									
1 (Oken 2010)	RCT	Not serious	Not serious	N/A	Not serious	10	10	SMD 1.27 (0.29, 2.25)	High

1. Crosses two lines of a defined MID
2. Crosses one line of a defined MID

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Usual care	Summary of results	

3. i^2>40%

G.11.1.12 Cranial electrotherapy stimulation

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Usual care	Summary of results	
Carer burden (lower values favour intervention)									
1 (Rose 2009)	RCT	Serious[1]	Serious[2]	N/A	Very serious[3]	19	19	SMD -0.14 (-0.78, 0.50)	Very low
Carer depression (lower values favour intervention)									
1 (Rose 2009)	RCT	Serious[1]	Serious[2]	N/A	Very serious[3]	19	19	SMD -0.38 (-1.02, 0.26)	Very low

1. Unclear reporting of methods
2. Not a relevant intervention in the UK
3. Crosses two lines of a defined MID

G.11.1.13 Psychotherapy versus psychoeducational interventions

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Psychotherapy	Psychoeducation	Summary of results	
Carer burden (lower values favour intervention)									
2	RCT	Not serious	Not serious	Not serious	Very serious[1]	30	30	SMD 0.16 (-0.34, 0.67)	Low
Carer depression (lower values favour intervention)									
3	RCT	Not serious	Not serious	Not serious	Serious[2]	63	64	SMD -0.29 (-0.64, 0.06)	Moderate
Carer anxiety (lower values favour intervention)									
1 (Gonyea 2016)	RCT	Serious[3]	Not serious	N/A	Very serious[1]	33	34	SMD -0.02 (-0.50, 0.46)	Very low
Carer self-efficacy (higher values favour intervention)									
1 (Gonyea 2016)	RCT	Serious[3]	Not serious	N/A	Very serious[1]	33	34	SMD 0.10 (-0.38, 0.58)	Very low
Behavioural and psychological symptoms of dementia – severity (lower values favour intervention)									
1 (Gonyea 2016)	RCT	Serious[3]	Not serious	N/A	Very serious[1]	33	34	SMD -0.20 (-0.68, 0.28)	Very low
Behavioural and psychological symptoms of dementia – reaction (lower values favour intervention)									
1 (Gonyea 2016)	RCT	Serious[3]	Not serious	N/A	Very serious[1]	33	34	SMD -0.26 (-0.74, 0.22)	Very low

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Psychotherapy	Psychoeducation	Summary of results	

1. Crosses two lines of a defined MID
2. Crosses one line of a defined MID
3. Unclear reporting of methods

G.11.1.14 CBT versus ACT (acceptance and commitment therapy)

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Usual care	Summary of results	
Carer depression (lower values favour intervention)									
1 (Losada 2015)	RCT	Serious[1]	Not serious	N/A	Serious[2]	42	45	SMD -0.27 (-0.69, 0.15)	Low
Carer anxiety (lower values favour intervention)									
1 (Losada 2015)	RCT	Serious[1]	Not serious	N/A	Very serious[3]	42	45	SMD -0.08 (-0.50, 0.34)	Very low

1. Unclear reporting of methods
2. Crosses one line of a defined MID
3. Crosses two lines of a defined MID

G.11.1.15 Spiritual care

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Usual care	Summary of results	
Carer self efficacy higher values favour intervention)									
1 (Salamizadeh 2016)	RCT	Very serious[1]	Not serious	N/A	Serious[2]	42	45	SMD 3.47 (0.60, 6.34)	Low

1. Unclear reporting of methods

G.11.1.16 Meta-regression

Quality assessment					Quality
Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	
73 (see appendix H for full list)	Not serious	Serious[1]	Not serious	Not serious	Moderate

1. Significant between study heterogeneity, with DICs suggesting more ompmlex models are not able to adequately resolve this heterogeneity

G.12 Staff training

G.12.1 Staff training

- What effect does training for staff working with people living with dementia have upon the experiences of people living with dementia in their care?

G.12.1.1 Residential care staff training: flexible education

Quality assessment						No of patients		Effect estimate	
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Usual care	Summary of results	Quality
Quality of life (self-rated) using QOL-AD (higher values favour intervention)									
1 (Beer 2011)	RCT	Not serious	Not serious	N/A	Serious[1]	161	190	MD 0.97 (-1.55, 3.49)	Moderate
Quality of life (carer-rated) using QOL-AD (higher values favour intervention)									
1 (Beer 2011)	RCT	Not serious	Not serious	N/A	Serious[1]	161	190	MD -1.07 (-3.34, 1.20)	Moderate
Quality of life (carer-rated) using ADRQOL (higher values favour intervention)									
1 (Beer 2011)	RCT	Not serious	Not serious	N/A	Serious[1]	161	190	MD -1.92 (-6.15, 2.31)	Moderate
Pain observed (Brief Pain Inventory) (higher values favour control)									
1 (Beer 2011)	RCT	Not serious	Not serious	N/A	Serious[2]	161	190	OR 1.98 (0.81, 4.83)	Moderate
Behavioural and psychological symptoms of dementia (NPI) (higher values favour control)									
1 (Beer 2011)	RCT	Not serious	Not serious	N/A	Very serious[3]	161	190	OR 1.18 (0.56, 2.49)	Low
Use of physical restraint observed (higher values favour control)									
1 (Beer 2011)	RCT	Not serious	Not serious	N/A	Very serious[3]	161	190	OR 1.06 (0.39, 2.91)	Low

1. Non-significant result
2. 95% CI crosses one line of a defined MID interval
3. 95% CI crosses two lines of a defined MID interval

G.12.1.2 Residential care staff training: activity provision

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Usual care	Summary of results	
Quality of life (QOL-AD) (higher values favour intervention)									
1 (Wenborn 2013)	RCT	Not serious	Not serious	N/A	Serious[1]	79	80	MD 0.26 (-3.04, 3.56)	Moderate
Cognition (MMSE) (higher values favour intervention)									
1 (Wenborn 2013)	RCT	Not serious	Not serious	N/A	Serious[1]	79	80	MD -0.36 (-2.22, 1.51)	Moderate
Behaviour and functional ability (CAPE-BRS) (higher values favour control)									
1 (Wenborn 2013)	RCT	Not serious	Not serious	N/A	Serious[1]	79	80	MD 0.52 (-1.63, 2.67)	Moderate
Challenging Behaviour Scale (higher values favour control)									
1 (Wenborn 2013)	RCT	Not serious	Not serious	N/A	Serious[1]	79	80	MD 4.13 (-21.10, 29.36)	Moderate
Cornell Scale for Depression in Dementia (higher values favour control)									
1 (Wenborn 2013)	RCT	Not serious	Not serious	N/A	Serious[1]	79	80	MD -0.09 (-1.33, 1.16)	Moderate
Rating Anxiety in Dementia (higher values favour control)									
1 (Wenborn 2013)	RCT	Not serious	Not serious	N/A	Serious[1]	79	80	MD 0.57 (-1.52, 2.66)	Moderate
Total number of medications (higher values favour control)									
1 (Wenborn 2013)	RCT	Not serious	Not serious	N/A	Serious[1]	79	80	MD -0.15 (-0.55, 0.24)	Moderate

1. Non-significant result

G.12.1.3 Residential care staff training: multi-sensory stimulation (snoezelen)

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Usual care	Summary of results	
Frequency of residents' smiling during the morning (higher values favour intervention)									
1 (van Weert 2005)	RCT	Serious[1]	Not serious	N/A	Not serious	60	61	MD 2.87 (0.93, 4.81)	Moderate
Change in residents' verbal communication - affective (positive) (estimated number of utterances per category) (higher values favour intervention)									
1 (van Weert 2005)	RCT	Serious[1]	Not serious	N/A	Not serious	60	61	MD 19.15 (9.31, 28.99)	Moderate
Change in residents' verbal communication - affective (negative) (estimated number of utterances per category) (higher values favour control)									
1 (van Weert 2005)	RCT	Serious[1]	Not serious	N/A	Not serious	60	61	MD -1.75 (-2.58, -0.92)	Moderate
Change in residents' verbal communication - instrumental (positive) (estimated number of utterances per category) (higher values favour intervention)									
1 (van Weert 2005)	RCT	Serious[1]	Not serious	N/A	Not serious	60	61	MD 38.40 (25.51, 51.29)	Moderate
Change in residents' verbal communication - instrumental (negative) (estimated number of utterances per category) (higher values favour control)									
1 (van Weert 2005)	RCT	Serious[1]	Not serious	N/A	Not serious	60	61	MD -2.02 (-3.41, -0.63)	Moderate
Mean duration of morning care (minutes) (higher values favour control)									
1 (van Weert 2005)	RCT	Serious[1]	Not serious	N/A	Not serious	60	61	MD 3.98 (1.76, 6.20)	Moderate

1. High dropout rates during study

G.12.1.4 Residential care staff training: behaviour management with motivational system

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Usual care	Summary of results	
Resident agitation (% of time) (higher values favour control)									
1 (Burgio 2002)	RCT	Very serious[1]	Not serious	N/A	Serious[2]	47	32	MD 0.60 (-4.81, 6.01)	Very low

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Usual care	Summary of results	
1. Potential contamination of the control group as they were also provided with training; unclear method of randomisation 2. Non-significant result									

G.12.1.5 Residential care staff training: feeding skills

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Usual care	Summary of results	
Food intake (higher values favour intervention)									
1 (Chang 2005)	RCT	Very serious[1]	Not serious	N/A	Serious[2]	12	8	MD -0.21 (-0.40, -0.02)	Very low
Edinburgh Feeding Evaluation in Dementia (higher values favour control)									
1 (Chang 2005)	RCT	Very serious[1]	Not serious	N/A	Serious[2]	12	8	MD 2.70 (0.66, 4.74)	Very low
1. Study at high risk of bias 2. Small sample size makes it difficult to have confidence in the effect estimates									

G.12.1.6 Residential care staff training: dementia care mapping

No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Usual care	Summary of results	
Agitation (CMAI) (higher values favour control)									
1 (Chenoweth 2009)	RCT	Not serious	Not serious	N/A	Not serious	95	64	MD -10.90 (-21.10, 0.70)	High
Neuropsychiatric inventory (higher values favour control)									
1 (Chenoweth 2009)	RCT	Not serious	Not serious	N/A	Serious[1]	95	64	MD 2.40 (-12.02, 16.82)	Moderate
Quality of life (QUALID) (higher values favour control)									

1 (Chenoweth 2009)	RCT	Not serious	Not serious	N/A	Serious[1]	95	64	MD -0.20 (-4.78, 4.38)	Moderate
Falls (higher values favour control)									
1 (Chenoweth 2009)	RCT	Not serious	Not serious	N/A	Not serious	95	64	MD -0.24 (-0.40, -0.08)	High

1. Non-significant result

G.12.1.7 Residential care staff training: person-centred care

Quality assessment						No of patients		Effect estimate	Quality
No of studies	**Design**	**Risk of bias**	**Indirectness**	**Inconsistency**	**Imprecision**	**Intervention**	**Usual care**	**Summary of results**	
Agitation (CMAI) (higher values favour control)									
2 (Chenoweth 2009, Chenoweth 2014)	RCT	Not serious	Not serious	Not serious	Not serious	141	128	MD -14.78 (-23.11, -6.45)	High
Neuropsychiatric inventory (higher values favour control)									
1 (Chenoweth 2009)	RCT	Not serious	Not serious	NA	Not serious	77	64	MD -7.10 (-9.12, -5.08)	High
Quality of life (QUALID and DemQOL) (higher values favour control)									
2 (Chenoweth 2009, Chenoweth 2014)	RCT	Not serious	Not serious	Not serious	Serious[1]	141	128	SMD -0.26 (-0.50, -0.02)	Moderate
Falls (higher values favour control)									
1 (Chenoweth 2009)	RCT	Not serious	Not serious	N/A	Not serious	77	64	MD -0.15 (-0.28, -0.02)	High

1. Crosses one line of a defined minimally important difference

G.12.1.8 Residential care staff training: awareness and communication

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Usual care	Summary of results	
Quality of life (QUALID - measured by family member) (higher values favour control)									
1 (Clare 2013)	RCT	Serious[1]	Not serious	N/A	Not serious	32	33	MD -3.98 (-7.60, -0.36)	Moderate
Well-being (Positive Response Schedule) (higher values favour intervention)									
1 (Clare 2013)	RCT	Serious[1]	Not serious	N/A	Serious[2]	32	33	MD 2.68 (-3.55, 8.91)	Low
Cognitive function (GADS) (higher values favour control) (higher values favour intervention)									
1 (Clare 2013)	RCT	Serious[1]	Not serious	N/A	Serious[2]	32	33	MD -1.18 (-3.44, 1.08)	Low
Behaviour - self-care (Behavioural Assessment Scale of Later Life) (higher values favour intervention)									
1 (Clare 2013)	RCT	Serious[1]	Not serious	N/A	Serious[2]	32	33	MD 0.56 (-1.06, 2.18)	Low
Behaviour - sensory abilities (BASOLL) (higher values favour intervention)									
1 (Clare 2013)	RCT	Serious[1]	Not serious	N/A	Serious[2]	32	33	MD -0.04 (-0.51, 0.43)	Low
Behaviour - mobility (BASOLL) (higher values favour intervention)									
1 (Clare 2013)	RCT	Serious[1]	Not serious	N/A	Serious[2]	32	33	MD -0.18 (-0.47, 0.11)	Low

1. Randomisation by care home, with only a small number of homes in the study
2. Non-significant result

G.12.1.9 Residential care staff training: challenging behaviours

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Usual care	Summary of results	
Agitation (CMAI) (higher values favour control)									

Quality assessment						No of patients		Effect estimate	Quality
No of studies	**Design**	**Risk of bias**	**Indirectness**	**Inconsistency**	**Imprecision**	**Intervention**	**Usual care**	**Summary of results**	
2 (Davison 2007, Deudon 2009)	RCT	Serious[1]	Not serious	Not serious	Not serious	204	146	MD -5.42 (-9.34, -1.50)	Moderate
Physically aggressive behaviour (higher values favour control)									
2 (Deudon 2009, Visser 2008)	RCT	Serious[2]	Not serious	Not serious	Serious[4]	179	146	SMD -0.03 (-0.25, 0.19)	Low
Verbally aggressive behaviour (higher values favour control)									
2 (Deudon 2009, Visser 2008)	RCT	Serious[2]	Not serious	Serious[7]	Very serious[6]	179	146	SMD 0.02 (-0.59, 0.63)	Very low
Quality of life (higher values favour intervention)									
1 (Deudon 2009)	RCT	Not serious	Not serious	N/A	Serious[5]	158	114	MD 1.51 (-0.41, 3.43)	Moderate
Quality of life (social interaction) (higher values favour control)									
1 (Visser 2008)	RCT	Very serious[3]	Not serious	N/A	Serious[5]	21	32	MD -5.36 (-15.69, 4.97)	Very low
Quality of life (feeling and mood) (higher values favour intervention)									
1 (Visser 2008)	RCT	Very serious[3]	Not serious	N/A	Serious[5]	21	32	MD 2.22 (-7.94, 12.38)	Very low
Quality of life (enjoyment of activities) (higher values favour intervention)									
1 (Visser 2008)	RCT	Very serious[3]	Not serious	N/A	Serious[5]	21	32	MD -4.90 (-24.68, 14.88)	Very low
Quality of life (awareness of self) (higher values favour intervention)									
1 (Visser 2008)	RCT	Very serious[3]	Not serious	N/A	Not serious	21	32	MD -15.79 (-31.40, -0.18)	Low
Mean number of hospitalisations (higher values favour control)									

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Usual care	Summary of results	
1 (Deudon 2009)	RCT	Not serious	Not serious	N/A	Very serious[6]	158	114	RR 0.63 (0.31, 1.26)	Low
Mean number of psychotropic drugs (higher values favour control)									
1 (Deudon 2009)	RCT	Not serious	Not serious	N/A	Serious[5]	158	114	MD -0.14 (-0.50, 0.22)	Moderate

1. High levels of attrition during study
2. Unclear reporting of one study in the meta-analysis
3. Unclear reporting of study
4. Crosses one line of a defined minimally important difference
5. Non-significant result
6. Crosses two lines of a defined minimally important difference
7. $i^2 > 40\%$

G.12.1.10 Residential care staff training: challenging behaviours with peer support

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Usual care	Summary of results	
Frequency of challenging behaviours (CMAI) (higher values favour control)									
1 (Davison 2007)	RCT	Serious[1]	Not serious	N/A	Serious[3]	35	32	MD -1.35 (-13.09, 10.39)	Low
Physically non-aggressive (higher values favour control)									
1 (Visser 2008)	RCT	Very serious[2]	Not serious	N/A	Serious[3]	23	32	MD 0.59 (-4.70, 5.88)	Very low
Physically aggressive (higher values favour control)									
1 (Visser 2008)	RCT	Very serious[2]	Not serious	N/A	Serious[3]	23	32	MD -1.85 (-9.56, 5.86)	Very low
Verbally non-aggressive (higher values favour control)									

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Usual care	Summary of results	
1 (Visser 2008)	RCT	Very serious[2]	Not serious	N/A	Serious[3]	23	32	MD 0.66 (-2.82, 4.14)	Very low
Verbally aggressive (higher values favour control)									
1 (Visser 2008)	RCT	Very serious[2]	Not serious	N/A	Serious[3]	23	32	MD 1.06 (-0.59, 2.71)	Very low
Quality of life (social interaction) (higher values favour intervention)									
1 (Visser 2008)	RCT	Very serious[2]	Not serious	N/A	Serious[3]	23	32	MD 4.40 (-6.83, 15.63)	Very low
Quality of life (awareness of self) (higher values favour intervention)									
1 (Visser 2008)	RCT	Very serious[2]	Not serious	N/A	Serious[3]	23	32	MD -2.60 (-18.82, 13.62)	Very low
Quality of life (feeling and mood) (higher values favour intervention)									
1 (Visser 2008)	RCT	Very serious[2]	Not serious	N/A	Not serious	23	32	MD 13.70 (3.50, 23.90)	Low
Quality of life (enjoyment of activities) (higher values favour intervention)									
1 (Visser 2008)	RCT	Very serious[2]	Not serious	N/A	Serious[1]	23	32	MD -8.48 (-25.60, 8.64)	Very low

1. High levels of attrition during study
2. Unclear reporting of study
3. Non-significant result

G.12.1.11 Residential care staff training: communication skills

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Usual care	Summary of results	
Cornell Scale for Depression in Dementia - mood related (higher values favour control)									
1 (McCallion 1999)	RCT	Serious[1]	Not serious	N/A	Not serious	49	56	MD -1.41 (-2.20, -0.62)	Moderate

Quality assessment						No of patients		Effect estimate	Quality
No of studies	**Design**	**Risk of bias**	**Indirectness**	**Inconsistency**	**Imprecision**	**Intervention**	**Usual care**	**Summary of results**	
Cornell Scale for Depression in Dementia - behavioural disturbance (higher values favour control)									
1 (McCallion 1999)	RCT	Serious[1]	Not serious	N/A	Not serious	49	56	MD -0.90 (-1.37, -0.43)	Moderate
Cornell Scale for Depression in Dementia - physical signs (higher values favour control)									
1 (McCallion 1999)	RCT	Serious[1]	Not serious	N/A	Not serious	49	56	MD -0.83 (-1.37, -0.29)	Moderate
Cornell Scale for Depression in Dementia - cyclic disturbance (higher values favour control)									
1 (McCallion 1999)	RCT	Serious[1]	Not serious	N/A	Not serious	49	56	MD -1.11 (-1.63, -0.59)	Moderate
Cornell Scale for Depression in Dementia - ideational disturbance (higher values favour control)									
1 (McCallion 1999)	RCT	Serious[1]	Not serious	N/A	Not serious	49	56	MD -0.51 (-0.82, -0.20)	Moderate
Cohen-Mansfield Agitation Inventory - aggressive behaviour (higher values favour control)									
1 (McCallion 1999)	RCT	Serious[1]	Not serious	N/A	Serious[2]	49	56	MD -1.72 (-4.56, 1.12)	Low
Cohen-Mansfield Agitation Inventory - physically nonaggressive behaviour (higher values favour control)									
1 (McCallion 1999)	RCT	Serious[1]	Not serious	N/A	Serious[2]	49	56	MD -0.40 (-2.76, 1.96)	Low
Cohen-Mansfield Agitation Inventory - verbally aggressive behaviour (higher values favour control)									
1 (McCallion 1999)	RCT	Serious[1]	Not serious	N/A	Not serious	49	56	MD -4.95 (-7.91, -1.99)	Moderate
Use of restraints – mechanical (higher values favour intervention)									
1 (McCallion 1999)	RCT	Serious[1]	Not serious	N/A	Not serious	49	56	MD 0.75 (0.12, 1.38)	Moderate
Use of restraints – chemical (higher values favour intervention)									
1 (McCallion 1999)	RCT	Serious[1]	Not serious	N/A	Serious[2]	49	56	MD 0.37 (-0.38, 1.12)	Low

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Usual care	Summary of results	
Multidimensional Observation Scale for Elderly Subjects – disorientation (higher values favour control)									
1 (McCallion 1999)	RCT	Serious[1]	Not serious	N/A	Not serious	49	56	MD 3.60 (0.70, 6.50)	Moderate
Multidimensional Observation Scale for Elderly Subjects – irritability (higher values favour control)									
1 (McCallion 1999)	RCT	Serious[1]	Not serious	N/A	Serious[2]	49	56	MD -1.68 (-3.96, 0.60)	Low
Multidimensional Observation Scale for Elderly Subjects – withdrawal (higher values favour control)									
1 (McCallion 1999)	RCT	Serious[1]	Not serious	N/A	Serious[2]	49	56	MD 0.21 (-1.50, 1.92)	Low

1. Method of randomisation and levels of loss to follow-up unclear
2. Non-significant result

G.12.1.12 Residential care staff training: emotion-oriented care

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Usual care	Summary of results	
Unstable affect (Cornell depression scale + BIP) (higher values favour control)									
1 (Finnema 2005)	RCT	Not serious	Not serious	N/A	Serious[1]	67	79	MD -0.87 (-2.02, 0.28)	Moderate
Cognitive adaption (BIP5 rebellious behaviour (0-15)) (higher values favour control)									
1 (Finnema 2005)	RCT	Not serious	Not serious	N/A	Serious[1]	67	79	MD -0.07 (-0.93, 0.79)	Moderate
Agitation (CMAI + BIP) (higher values favour control) (higher values favour intervention)									
1 (Finnema 2005)	RCT	Not serious	Not serious	N/A	Serious[1]	67	79	MD 0.78 (-0.34, 1.90)	Moderate
PGCMS dissatisfaction with present situation (0-4) (higher values favour intervention)									

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Usual care	Summary of results	
1 (Finnema 2005)	RCT	Not serious	Not serious	N/A	Serious[1]	67	79	MD 0.25 (-0.07, 0.57)	Moderate
PGCMS attitude towards ageing (0-6) (higher values favour intervention)									
1 (Finnema 2005)	RCT	Not serious	Not serious	N/A	Not serious	67	79	MD 0.80 (0.46, 1.14)	High
Developing and maintaining social relationships questionnaire (higher values favour intervention)									
1 (Finnema 2005)	RCT	Not serious	Not serious	N/A	Serious[1]	67	79	MD -0.50 (-1.73, 0.73)	Moderate
Coping with nursing home environment (BIP + ASEP4 inactivity + GRGS-other activity) (higher values favour intervention)									
1 (Finnema 2005)	RCT	Not serious	Not serious	N/A	Serious[1]	67	79	MD 0.24 (-0.95, 1.43)	Moderate

1. Non-significant result

G.12.1.13 Residential care staff training: reducing antipsychotic drug use

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Usual care	Summary of results	
Proportion taking neuroleptics (lower numbers favour intervention)									
1 (Fossey 2006)	RCT	Serious[1]	N/A	Not serious	Not serious	176	170	RR 0.55 (0.39, 0.76)	Moderate
Fall in past 12 months (lower numbers favour intervention)									
1 (Fossey 2006)	RCT	Serious[1]	N/A	Not serious	Very serious[3]	176	170	RR 0.90 (0.59, 1.38)	Very low
Aggression (Cohen-Mansfield agitation score - lower numbers favour intervention)									
1 (Fossey 2006)	RCT	Serious[1]	N/A	Not serious	Serious[2]	176	170	MD 0.3 (-8.3, 8.9)	Low
Wellbeing (dementia care mapping - higher numbers favour intervention)									

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Usual care	Summary of results	
1 (Fossey 2006)	RCT	Serious[1]	N/A	Not serious	Serious[2]	176	170	MD -0.2 (-0.5, 0.2)	Low

1. Lack of appropriate blinding
2. Non-significant result
3. 95% CI crosses two lines of a defined MID interval

G.12.1.14 Residential care staff training: towel bathing

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Usual care	Summary of results	
Any agitation or aggression (%time – higher numbers favour control)									
1 (Sloane 2004)	RCT	Serious[1]	Not serious	N/A	Serious[2]	24	24	MD -11.22 (-23.89, 1.45)	Low
Any physical agitation or aggression (%time – higher numbers favour control)									
1 (Sloane 2004)	RCT	Serious[1]	Not serious	N/A	Serious[2]	24	24	MD -0.59 (-1.30, 0.12)	Low
Any aggression (rate/15minutes – higher numbers favour control)									
1 (Sloane 2004)	RCT	Serious[1]	Not serious	N/A	Not serious	24	24	MD -1.08 (-1.86, -0.30)	Moderate
Hit, bite, kick, throw or spit (rate/15 minutes – higher numbers favour control)									
1 (Sloane 2004)	RCT	Serious[1]	Not serious	N/A	Serious[2]	24	24	MD -0.16 (-0.48, 0.16)	Low
Other aggression (attempts/grabbing, rate/15 minutes – higher numbers favour control)									
1 (Sloane 2004)	RCT	Serious[1]	Not serious	N/A	Not serious	24	24	MD -0.97 (-1.74, -0.20)	Moderate
Yelling, crying, moaning (%time – higher numbers favour control)									

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Usual care	Summary of results	
1 (Sloane 2004)	RCT	Serious[1]	Not serious	N/A	Serious[2]	24	24	MD -0.31 (-0.90, 0.28)	Low
Complaints, threats (rate/15 minutes – higher numbers favour control)									
1 (Sloane 2004)	RCT	Serious[1]	Not serious	N/A	Serious[2]	24	24	MD -0.72 (-1.71, 0.27)	Low
Mean discomfort scale score (higher numbers favour control)									
1 (Sloane 2004)	RCT	Serious[1]	Not serious	N/A	Not serious	24	24	MD -0.56 (-0.83, -0.29)	Moderate

1. No information on study dropouts
2. Non-significant result

G.12.1.15 Residential care staff training: person-centred showering

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Usual care	Summary of results	
Any agitation or aggression (%time – higher numbers favour control)									
1 (Sloane 2004)	RCT	Serious[1]	Not serious	N/A	Serious[2]	25	24	MD -8.89 (-23.38, 5.60)	Low
Any physical agitation or aggression (%time – higher numbers favour control)									
1 (Sloane 2004)	RCT	Serious[1]	Not serious	N/A	Serious[2]	25	24	MD -0.39 (-1.67, 0.89)	Low
Any aggression (rate/15minutes – higher numbers favour control)									
1 (Sloane 2004)	RCT	Serious[1]	Not serious	N/A	Not serious	25	24	MD -0.94 (-1.75, -0.13)	Moderate
Hit, bite, kick, throw or spit (rate/15 minutes – higher numbers favour control)									
1 (Sloane 2004)	RCT	Serious[1]	Not serious	N/A	Serious[2]	25	24	MD -0.33 (-0.73, 0.07)	Low

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Usual care	Summary of results	
Other aggression (attempts/grabbing, rate/15 minutes – higher numbers favour control)									
1 (Sloane 2004)	RCT	Serious[1]	Not serious	N/A	Not serious	25	24	MD -0.78 (-1.54, -0.02)	Moderate
Yelling, crying, moaning (%time – higher numbers favour control)									
1 (Sloane 2004)	RCT	Serious[1]	Not serious	N/A	Serious[2]	25	24	MD -0.09 (-0.69, 0.51)	Low
Complaints, threats (rate/15 minutes – higher numbers favour control)									
1 (Sloane 2004)	RCT	Serious[1]	Not serious	N/A	Serious[2]	25	24	MD -0.39 (-1.35, 0.57)	Low
Mean discomfort scale score (higher numbers favour control)									
1 (Sloane 2004)	RCT	Serious[1]	Not serious	N/A	Not serious	25	24	MD -0.31 (-0.54, -0.08)	Moderate

1. No information on study dropouts
2. Non-significant result

G.12.1.16 Residential care staff training: apathy management

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Usual care	Summary of results	
NPI – affect (higher numbers favour control)									
1 (Leone 2013)	RCT	Serious[1]	Not serious	N/A	Serious[2]	119	111	MD 0.91 (-0.63, 2.45)	Low
NPI – apathy (higher numbers favour control)									
1 (Leone 2013)	RCT	Serious[1]	Not serious	N/A	Serious[2]	119	111	MD 0.11 (-1.09, 1.31)	Low
NPI – hyperactivity (higher numbers favour control)									
1 (Leone 2013)	RCT	Serious[1]	Not serious	N/A	Serious[2]	119	111	MD 0.40 (-2.23, 3.03)	Low
NPI – psychotic symptoms (higher numbers favour control)									

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Usual care	Summary of results	
1 (Leone 2013)	RCT	Serious[1]	Not serious	N/A	Serious[2]	119	111	MD 0.60 (-0.70, 1.90)	Low
ADL Katz scale – toileting (higher numbers favour intervention)									
1 (Leone 2013)	RCT	Serious[1]	Not serious	N/A	Not serious	119	111	MD -0.18 (-0.29, -0.07)	Moderate
ADL Katz scale – dressing (higher numbers favour intervention)									
1 (Leone 2013)	RCT	Serious[1]	Not serious	N/A	Serious[2]	119	111	MD -0.08 (-0.27, 0.11)	Low
ADL Katz scale – going to the toilet (higher numbers favour intervention)									
1 (Leone 2013)	RCT	Serious[1]	Not serious	N/A	Serious[2]	119	111	MD 0.13 (-0.08, 0.34)	Low
ADL Katz scale – transferring (higher numbers favour intervention)									
1 (Leone 2013)	RCT	Serious[1]	Not serious	N/A	Serious[2]	119	111	MD -0.12 (-0.26, 0.02)	Low
ADL Katz scale – continence (higher numbers favour intervention)									
1 (Leone 2013)	RCT	Serious[1]	Not serious	N/A	Serious[2]	119	111	MD 0.16 (-0.02, 0.34)	Low
ADL Katz scale – feeding (higher numbers favour intervention)									
1 (Leone 2013)	RCT	Serious[1]	Not serious	N/A	Serious[2]	119	111	MD 0.05 (-0.16, 0.26)	Low
Apathy inventory – emotional blunting (higher numbers favour control)									
1 (Leone 2013)	RCT	Serious[1]	Not serious	N/A	Not serious	119	111	MD -0.50 (-0.84, -0.16)	Moderate
Apathy inventory – lack of initiative (higher numbers favour control)									
1 (Leone 2013)	RCT	Serious[1]	Not serious	N/A	Serious[2]	119	111	MD -0.20 (-0.47, 0.07)	Low
Apathy inventory – lack of interest (higher numbers favour control)									
1 (Leone 2013)	RCT	Serious[1]	Not serious	N/A	Serious[2]	119	111	MD 0.06 (-0.20, 0.32)	Low

1. Unclear method of randomisation
2. Non-significant result

G.12.1.17 Residential care staff training: sensitivity to non-verbal emotion signals

Quality assessment						No of patients		Effect estimate	Quality
No of studies	**Design**	**Risk of bias**	**Indirectness**	**Inconsistency**	**Imprecision**	**Intervention**	**Usual care**	**Summary of results**	
Symptomatology (higher numbers favour control)									
1 (Magai 2002)	RCT	Very serious[1]	Not serious	N/A	Serious[2]	34	23	MD -39.20 (-57.15, -21.25)	Very low
Positive emotion (higher numbers favour intervention)									
1 (Magai 2002)	RCT	Very serious[1]	Not serious	N/A	Serious[3]	41	27	MD 0.70 (-0.89, 2.29)	Very low
Negative emotion (higher numbers favour control)									
1 (Magai 2002)	RCT	Very serious[1]	Not serious	N/A	Serious[3]	41	27	MD 0.10 (-1.49, 1.69)	Very low
Brief symptom inventory (higher numbers favour control)									
1 (Magai 2002)	RCT	Very serious[1]	Not serious	N/A	Serious[3]	8	5	MD -4.90 (-14.34, 4.54)	Very low

1. Large differences in baseline characteristics between the intervention and control groups, including in outcome measures
2. Significant differences between the intervention and control groups at baseline in this outcome, which may be a confounding factor in the mean change data
3. Non-significant result

G.12.1.18 Residential care staff and nurse training: effective communication, empathy development and conflict resolution

Quality assessment						No of patients		Effect estimate	Quality
No of studies	**Design**	**Risk of bias**	**Indirectness**	**Inconsistency**	**Imprecision**	**Intervention**	**Usual care**	**Summary of results**	
Staff easy to talk to (higher numbers favour intervention)									
1 (Robison 2007)	RCT	Serious[1]	Not serious	N/A	Not serious	169	156	MD 0.19 (0.02, 0.36)	Moderate
Staff behaviours scale (higher numbers favour intervention)									
1 (Robison 2007)	RCT	Serious[1]	Not serious	N/A	Not serious	169	156	MD 0.67 (0.11, 1.23)	Moderate

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Usual care	Summary of results	
Family involvement scale - spouses (higher numbers favour intervention)									
1 (Robison 2007)	RCT	Serious[1]	Not serious	N/A	Serious[2]	169	156	MD 0.96 (-0.54, 2.46)	Low
Family involvement scale – adult children (higher numbers favour intervention)									
1 (Robison 2007)	RCT	Serious[1]	Not serious	N/A	Serious[2]	169	156	MD 0.28 (-0.34, 0.90)	Low

1. Method of randomisation unclear
2. Non-significant result

G.12.1.19 Residential care staff and nurse training: restraint use reduction

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Usual care	Summary of results	
Proportion of residents restrained (higher values favour control)									
1 (Pellfolk 2010)	RCT	Serious[1]	Not serious	N/A	Not serious	149	139	RR 0.53 (0.36, 0.77)	Moderate
Frequency of use of physical restraints (higher numbers favour control)									
1 (Testad 2005)	RCT	Very serious[2]	Not serious	N/A	Not serious	55	87	MD -2.40 (-4.35, -0.45)	Low
Proportion of residents prescribed neuroleptics (higher numbers favour control)									
1 (Pellfolk 2010)	RCT	Serious[1]	Not serious	N/A	Serious[4]	144	127	RR 1.24 (0.94, 1.64)	Low
Proportion of residents experiencing paralysis (higher numbers favour control)									
1 (Pellfolk 2010)	RCT	Serious[1]	Not serious	N/A	Very serious[5]	138	127	RR 1.07 (0.66, 1.72)	Very low
Proportion of residents walking independently (higher numbers favour intervention)									
1 (Pellfolk 2010)	RCT	Serious[1]	Not serious	N/A	Serious[4]	142	129	RR 1.16 (0.93, 1.46)	Low

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Usual care	Summary of results	
Proportion of residents able to rise from their bed (higher numbers favour intervention)									
1 (Pellfolk 2010)	RCT	Serious[1]	Not serious	N/A	Serious[4]	141	129	RR 1.04 (0.87, 1.25)	Low
Proportion of residents able to rise from a chair (higher numbers favour intervention)									
1 (Pellfolk 2010)	RCT	Serious[1]	Not serious	N/A	Serious[4]	142	128	RR 1.13 (0.96, 1.32)	Low
Proportion of residents needing an aid when walking (higher numbers favour control)									
1 (Pellfolk 2010)	RCT	Serious[1]	Not serious	N/A	Serious[4]	140	124	RR 1.11 (0.91, 1.34)	Low
Staff assessment of fall risk (higher numbers favour control)									
1 (Pellfolk 2010)	RCT	Serious[1]	Not serious	N/A	Serious[3]	140	120	MD -2.90 (-10.64, 4.84)	Low
Proportion of people falling (higher numbers favour control									
1 (Pellfolk 2010)	RCT	Serious[1]	Not serious	N/A	Very serious[5]	149	139	RR 1.17 (0.57, 2.40)	Very low
Agitation (higher numbers favour control)									
2 (Testad 2005, Testad 2010)	RCT	Very serious[2]	Not serious	Serious[6]	Very serious[5]	99	133	SMD -0.08 (-0.90, 0.75)	Very low
Proportion of residents who hit others (higher numbers favour control)									
1 (Pellfolk 2010)	RCT	Serious[1]	Not serious	N/A	Very serious[5]	141	130	RR 1.23 (0.79, 1.91)	Very low
Proportion of residents who make aggressive threats (higher numbers favour control)									
1 (Pellfolk 2010)	RCT	Serious[1]	Not serious	N/A	Serious[4]	142	131	RR 0.91 (0.70, 1.18)	Low
Proportion of residents with wandering behaviour (higher numbers favour control)									
1 (Pellfolk 2010)	RCT	Serious[1]	Not serious	N/A	Serious[4]	142	131	RR 1.24 (0.91, 1.69)	Low

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Usual care	Summary of results	

1. High level of attrition in study
2. Major differences in baseline characteristics between the two arms of the trial
3. Non-significant result
4. 95% CI crosses one line of a defined MID interval
5. 95% CI crosses two lines of a defined MID interval
6. $i^2 > 40\%$

G.12.1.20 Residential care nurse training: managing depression nursing guideline

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Usual care	Summary of results	
Depression (MDS/RAI-DRS – higher numbers favour control)									
1 (Verkaik 2011)	RCT	Not serious	Not serious	N/A	Serious[1]	62	35	MD -1.00 (-2.41, 0.41)	Moderate
Depression (Cornell Scale – higher numbers favour control)									
1 (Verkaik 2011)	RCT	Not serious	Not serious	N/A	Serious[1]	62	35	MD 0.09 (-2.56, 2.74)	Moderate
Mood (morning care – higher numbers favour intervention)									
1 (Verkaik 2011)	RCT	Not serious	Not serious	N/A	Serious[1]	62	35	MD -0.01 (-0.34, 0.32)	Moderate
Mood (living room – higher numbers favour intervention)									
1 (Verkaik 2011)	RCT	Not serious	Not serious	N/A	Serious[1]	62	35	MD -0.09 (-0.35, 0.17)	Moderate

1. Non-significant result

G.12.1.21 Residential care nurse training: restraint reduction

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Usual care	Summary of results	
Mean restraint intensity (higher numbers favour control)									
1 (Huizing 2006)	RCT	Serious[1]	Not serious	N/A	Serious[2]	72	54	MD -0.35 (-0.96, 0.26)	Low

1. Method of randomisation not specified
2. Non-significant result

G.12.1.22 Residential care nurse training: dementia care mapping

No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Usual care	Summary of results	
Agitation (CMAI – higher numbers favour control)									
1 (van de Ven 2013)	RCT	Serious[1]	Not serious	N/A	Serious[2]	73	119	MD 1.05 (-4.89, 6.99)	Low
Behavioural symptoms (NPI-NH – higher numbers favour control)									
1 (van de Ven 2013)	RCT	Serious[1]	Not serious	N/A	Not serious	73	119	MD 3.08 (0.61, 5.55)	Moderate
Quality of life (Qualidem - higher numbers favour intervention)									
1 (van de Ven 2013)	RCT	Serious[1]	Not serious	N/A	Serious[2]	73	119	MD 0.13 (-5.53, 5.79)	Low
Quality of life (EQ-5D - higher numbers favour intervention)									
1 (van de Ven 2013)	RCT	Serious[1]	Not serious	N/A	Serious[2]	73	119	MD 0.04 (-0.03, 0.11)	Low

1. Method of randomisation not specified
2. Non-significant result

G.12.1.23 Occupational therapist training: interdisciplinary training

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Usual care	Summary of results	
AMPS process (higher numbers favour control)									
1 (Döpp 2015)	RCT	Not serious	Not serious	N/A	Very serious[1]	21	12	MD 0.20 (-0.11, 0.51)	Low
AMPS motor (higher numbers favour control)									
1 (Döpp 2015)	RCT	Not serious	Not serious	N/A	Very serious[1]	21	12	MD 0.30 (-0.05, 0.65)	Low
Interview for Deterioration of Daily Activities in Dementia (higher numbers favour control)									
1 (Döpp 2015)	RCT	Not serious	Not serious	N/A	Very serious[1]	21	12	MD -0.30 (-5.72, 5.12)	Low
Canadian Occupational Performance Measure – performance (higher numbers favour intervention)									
1 (Döpp 2015)	RCT	Not serious	Not serious	N/A	Very serious[1]	21	12	MD -0.30 (-1.53, 0.93)	Low
Canadian Occupational Performance Measure – satisfaction (higher numbers favour intervention)									
1 (Döpp 2015)	RCT	Not serious	Not serious	N/A	Very serious[1]	21	12	MD 0.40 (-0.81, 1.61)	Low
DQOL – overall (higher numbers favour intervention)									
1 (Döpp 2015)	RCT	Not serious	Not serious	N/A	Very serious[1]	21	12	MD -0.40 (-0.95, 0.15)	Low
DQOL – aesthetics (higher numbers favour intervention)									
1 (Döpp 2015)	RCT	Not serious	Not serious	N/A	Very serious[1]	21	12	MD -3.20 (-6.50, 0.10)	Low
DQOL – positive affect (higher numbers favour intervention)									
1 (Döpp 2015)	RCT	Not serious	Not serious	N/A	Very serious[1]	21	12	MD 1.40 (-1.10, 3.90)	Low
DQOL – negative affect (higher numbers favour control)									

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Usual care	Summary of results	
1 (Döpp 2015)	RCT	Not serious	Not serious	N/A	Very serious[1]	21	12	MD -0.70 (-4.15, 2.75)	Low
DQOL – self-esteem (higher numbers favour intervention)									
1 (Döpp 2015)	RCT	Not serious	Not serious	N/A	Very serious[1]	21	12	MD 1.10 (-0.61, 2.81)	Low
DQOL – feelings of belonging (higher numbers favour intervention)									
1 (Döpp 2015)	RCT	Not serious	Not serious	N/A	Serious[2]	21	12	MD 1.30 (0.24, 2.36)	Moderate
EQ-5D (higher numbers favour intervention)									
1 (Döpp 2015)	RCT	Not serious	Not serious	N/A	Very serious[1]	21	12	MD -0.10 (-0.24, 0.04)	Low

1. Small sample size and non-significant result
2. Small sample size

G.12.1.24 GP training: flexible education

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Usual care	Summary of results	
Quality of life (self-rated) using QOL-AD (higher values favour intervention)									
1 (Beer 2011)	RCT	Not serious	Not serious	N/A	Serious[1]	157	194	MD -0.61 (-3.07, 1.85)	Moderate
Quality of life (carer-rated) using QOL-AD (higher values favour intervention)									
1 (Beer 2011)	RCT	Not serious	Not serious	N/A	Serious[1]	157	194	MD -0.07 (-2.31, 2.17)	Moderate
Quality of life (carer-rated) using ADRQOL (higher values favour intervention)									
1 (Beer 2011)	RCT	Not serious	Not serious	N/A	Serious[1]	157	194	MD 1.02 (-3.23, 5.27)	Moderate
Pain observed (Brief Pain Inventory) (log odds ratio) (higher values favour control)									

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Usual care	Summary of results	
1 (Beer 2011)	RCT	Not serious	Not serious	N/A	Very serious[2]	157	194	OR 0.60 (0.25, 1.47)	Low
Behavioural and psychological symptoms of dementia (NPI) (higher values favour control)									
1 (Beer 2011)	RCT	Not serious	Not serious	N/A	Very serious[2]	157	194	OR 0.81 (0.40, 1.61)	Low
Use of physical restraint observed (higher values favour control)									
1 (Beer 2011)	RCT	Not serious	Not serious	N/A	Serious[3]	157	194	OR 0.44 (0.17, 1.11)	Moderate

1. Non-significant result
2. 95% CI crosses two lines of a defined MID
3. 95% CI crosses one line of a defined MID

G.12.1.25 Pooled analysis: person-centred care versus control

Quality assessment						No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Intervention	Usual care	Summary of results	
Agitation using CMAI (higher values favour control)									
5 (Chenoweth 2009, Chenoweth 2014, Davison 2007, Deudon 2009, van de Ven 2013)	RCT	Not serious	Not serious	Not serious	Not serious	548	393	MD -4.70 (-7.75, -1.65)	High
NPI (higher numbers favour control)									
2 (Chenoweth 2009, van de Ven 2013)	RCT	Not serious	Not serious	Serious[1]	Serious[2]	245	183	MD -1.31 (-10.23, 7.61)	Low
Quality of life (higher numbers favour intervention)									

Quality assessment						No of patients		Effect estimate	Quality
No of studies	**Design**	**Risk of bias**	**Indirectness**	**Inconsistency**	**Imprecision**	**Intervention**	**Usual care**	**Summary of results**	
4 (Chenoweth 2009, Chenoweth 2014, Deudon 2009, van de Ven 2013)	RCT	Not serious	Not serious	Not serious	Serious[3]	467	361	SMD 0.15 (0.01, 0.29)	Moderate

1. $I^2 > 50\%$
2. Non-significant result
3. Crosses one line of a defined minimally important difference

G.13 Needs of younger people living with dementia

G.13.1 The specific needs of younger people living with dementia

Review question

- What are the specific needs of younger people living with dementia?

G.13.1.1 CERQual tables

Themes identified for employment: experiences and coping

Studies	Study design	Description	Methodological limitations	Relevance	Coherence	Adequacy	Confidence
Theme: PWD: An awareness of changes in their functioning in the work place as they developed dementia.							
1 (Chaplin 2016)	Interviews	For three participants, the Engineer, the Businessman and the Schools Meals Assistant, the first signs were poor short-term memory and a difficulty in remembering names and adjusting to new tasks.	Not serious	High	High	Low[1]	Low
Theme: PWD: Shock at losing their expected future.							
1 (Clemerson 2014)	Semi-structured interviews	For many, this included loss of employment as they were forced to take early retirement.	Not serious	High	High	Low[1]	Low
Theme: PWD: A reluctance to acknowledge the signs							
1 (Chaplin 2016)	Interviews	All of the participants described how they did not initially think that these difficulties in specific areas of functioning were the first signs of something more serious. At this stage, they tended to ascribe the changes to pressure of work, new work roles, life-long traits, such as poor memory or declining physical skills such as poor eyesight	Not serious	High	High	Low[1]	Low
Theme: PWD: Sharing the fears							

Studies	Study design	Description	Methodological limitations	Relevance	Coherence	Adequacy	Confidence
1 (Chaplin 2016)	Interviews	They then began to suspect it was something more serious and all discussed their difficulties with their partners and were encouraged to seek further help.	Not serious	High	High	Low[1]	Low
Theme: PWD: Self-management							
1 (Chaplin 2016)	Interviews	Three of the participants were able to discuss strategies for managing the symptoms of their illness in the workplace. They all spent more time and effort in planning and organising tasks and acknowledged how difficult it could be even with these strategies in place	Not serious	High	High	Low[1]	Low
Theme: PWD: Feeling under scrutiny							
1 (Chaplin 2016)	Interviews	The three participants who worked more closely with others described how their managers or colleagues had noticed that they were having difficulties in some tasks. They mainly tried to manage this by increased observation of the employee but did not discuss this with the employee. Consequently, the participants felt that they were being watched covertly and they would have preferred to have been consulted about this.	Not serious	High	High	Low[1]	Low
Theme: PWD: A lack of consultation about management decisions							
1 (Chaplin 2016)	Interviews	Though two of the participants were given some adjusted duties when their employers became aware that they were having difficulties, none of the participants said that they were offered any 'reasonable adjustments' to their work role under the Equality Act (2010) after diagnosis. None of the participants were referred to a Disability Employment Advisor by their workplace. The HGV Driver and the School Meals Assistant were advised to take sickness leave when their employers became aware of the extent of their difficulties at work. They were advised to seek further assessment of their difficulties from their GP. Both of their GP's did make referrals on, one to a Neurologist and one to a Psychiatrist. Both these participants were then on	Not serious	High	High	Low[1]	Low

Studies	Study design	Description	Methodological limitations	Relevance	Coherence	Adequacy	Confidence
		sickness leave for the full six months and never returned to work					
Theme: PWD: A belief in continued competence despite the realisation of impairment							
1 (Chaplin 2016)	Interviews	Three of the participants felt that they would have been able to carry on with an adjusted work role when they were diagnosed with dementia, while the School meals Assistant and the Businessman believed that they were no longer competent.	Not serious	High	High	Low[1]	Low
Theme: PWD: Feeling abandoned by the workplace and consequent feelings of resentment towards the workplace							
1 (Chaplin 2016)	Interviews	Three of the participants expressed feelings of abandonment in how their employment situation was managed by their workplace. They felt that when they received their diagnosis and informed their workplace, no real attempt was made to find any adjusted work role for them.	Not serious	High	High	Low[1]	Low
Theme: PWD: An acceptance of the final outcome							
1 (Chaplin 2016)	Interviews	Four of the participants expressed an acceptance of the final outcome of their employment	Not serious	High	High	Low[1]	Low
Theme: PWD: Coming to terms with their situation							
1 (Chaplin 2016)	Interviews	Two of the participants are now on Employment Support Allowance, one has taken early retirement and two classed themselves as semi-retired. Four of the participants said that their work was a big part of their life and that they had enjoyed it and taken a pride in doing it well.	Not serious	High	High	Low[1]	Low
Theme: PWD: Financial hardship and consequent worry							
1 (Chaplin 2016)	Interviews	All of the participants said that leaving work had affected their family and their relationships. The Nursing Assistant and the HGV Driver both had partners who are still working and they had taken on more domestic roles to help them. For the HGV Driver and the School Meals	Not serious	High	High	Low[1]	Low

Studies	Study design	Description	Methodological limitations	Relevance	Coherence	Adequacy	Confidence
		Assistant, leaving work had meant some financial hardship and consequent worry					
Theme: PWD: A positive outlook for the future							
1 (Chaplin 2016)	Interviews	Despite their difficult experiences all of the participants were determined to be positive about their future. All of the participants said that they had taken up new hobbies or restarted old ones since leaving or reducing their work. The three participants who are under the age of 65 had been referred to the Young Onset Dementia Service in their local area and had become involved in the various social and leisure activities facilitated by this service.	Not serious	High	High	Low[1]	Low

1. This is the only UK study that addresses this theme, and contains only a very small numbers of participants.

Themes identified for general experiences and coping

Studies	Study design	Description	Methodological limitations	Relevance	Coherence	Adequacy	Confidence
Theme: PWD: Relief at getting the diagnosis confirmed							
1 (Clayton-Turner 2015)	Interviews	Relief at getting the diagnosis confirmed	Serious[1]	High	High	Moderate[1]	Low
Theme: PWD: Feelings of shock and a sense of loss at receiving the diagnosis							
1 (Pipon-Young 2012)	Interviews, group discussions	Feelings of shock and a sense of loss at receiving the diagnosis	Not serious	High	High	Low[3]	Low
Theme: PWD: Experiences of feeling 'too young'.							
2 (Clemerson 2014, Pipon-	Semi-structured interviews, interviews,	What surprised people was their age at diagnosis, with the general assumption that dementia was something affecting older people.	Not serious	High	High	High	High

Studies	Study design	Description	Methodological limitations	Relevance	Coherence	Adequacy	Confidence
Young 2012)	group discussions						
Theme: PWD: Ambiguity of the term 'younger people with dementia'							
1 (Pipon-Young 2012)	Interviews, group discussions	Ambiguity of the term 'younger people with dementia', and people being unsure whether the label applied to them	Not serious	High	High	Low[3]	Low
Theme: PWD: Younger people living with dementia often have responsibility for children, a mortgage or a business to run							
1 (Pipon-Young 2012)	Interviews, group discussions	Younger people living with dementia often have responsibility for children, a mortgage or a business to run	Not serious	High	High	Low[3]	Low
Theme: PWD: People coped by normalising the situation.							
1 (Clemerson 2014)	Semi-structured interviews	Creating an identity as an older person, even transiently, allowed people to make sense of developing AD by normalising the life-cycle.	Serious[1]	High	High	Low[3]	Very low
Theme: PWD: Telling children about the diagnosis is difficult							
1 (Clayton-Turner 2015)	Interviews	Telling children about the diagnosis is difficult, particularly at an age when they will not have been expecting it	Serious[1]	High	High	Moderate[1]	Low
Theme: PWD: Developing dementia forced people to contemplate death.							
1 (Clemerson 2014)	Semi-structured interviews	Developing dementia forced people to contemplate death	Serious[1]	High	High	Low[3]	Very low
Theme: PWD: Shock at losing their expected future.							
1 (Clemerson 2014)	Semi-structured interviews	For many, this included loss of employment as they were forced to take early retirement	Serious[1]	High	High	Low[3]	Very low
Theme: PWD: Loss of adult competency.							

Studies	Study design	Description	Methodological limitations	Relevance	Coherence	Adequacy	Confidence
1 (Clemerson 2014)	Semi-structured interviews	Loss of adult competency represents another sub-theme in the disruption to the life-cycle. This emerged through people's experience of either feeling more 'childlike' due to a loss of skills or being treated this way by others	Serious[1]	High	High	Low[3]	Very low
Theme: PWD: Some people tried to prevent themselves from thinking about the future.							
1 (Clemerson 2014)	Semi-structured interviews	Some people tried to prevent themselves from thinking about the future	Serious[1]	High	High	Low[3]	Very low
Theme: PWD: Some people tried to stay positive, which for a few meant denying further significant decline.							
1 (Clemerson 2014)	Semi-structured interviews	Some people tried to stay positive, which for a few meant denying further significant decline	Serious[1]	High	High	Low[3]	Very low
Theme: PWD: With further reflection it seemed that some participants were working towards resolving concerns through comparing their situation to others who were more impaired or died younger than themselves.							
1 (Clemerson 2014)	Semi-structured interviews	With further reflection it seemed that some participants were working towards resolving concerns through comparing their situation to others who were more impaired or died younger than themselves.	Serious[1]	High	High	Low[3]	Very low
Theme: PWD: Redefining self							
2 (Clemerson 2014, Pipon-Young 2012)	Semi-structured interviews, interviews, group discussions	Acknowledging change. Descriptions of the experience of dementia often related to changes people experienced, particularly in relation to what they could no longer do, a loss of independence or how their life had changed. This included a loss in social status and an inability to carry out everyday tasks.	Not serious	High	High	High	High
Theme: PWD: All participants referred to their concerns of what may happen as their dementia progresses. This concern arose in response to meeting others with more advanced dementia.							
1 (Pipon-Young 2012)	Interviews, group discussions	This concern arose in response to meeting others with more advanced dementia. It was also frightening for	Not serious	High	High	Low[3]	Low

Studies	Study design	Description	Methodological limitations	Relevance	Coherence	Adequacy	Confidence
		people to imagine a time when they may not realize their memory was deteriorating.					
Theme: PWD: Often raised was the negative impact of others' perceptions.							
1 (Pipon-Young 2012)	Interviews, group discussions	Typically described were the negative perceptions of the word 'dementia', resulting in a lack of understanding about dementia and a loss as to how to be with people with dementia. A number of misconceptions were described regarding others' understanding of dementia. There seemed to be a sense that there was an avoidance of a true understanding in order to prevent painful truths.	Not serious	High	High	Low[3]	Low
Theme: PWD: A reduced sense of self-worth also contributed to the threat to self.							
1 (Clemerson 2014)	Semi-structured interviews	Simply having the disease made some individuals question their worth.	Serious[1]	High	High	Low[3]	Very low
Theme: PWD: Most participants who disclosed their condition had positive responses from others, which helped them to accept their diagnosis as part of who they were.							
1 (Clemerson 2014)	Semi-structured interviews	Most participants who disclosed their condition had positive responses from others, which helped them to accept their diagnosis as part of who they were.	Serious[1]	High	High	Low[3]	Very low
Theme: PWD: Holding on to their existing self-concept.							
2 (Clemerson 2014, Pipon-Young 2012)	Semi-structured interviews, interviews, group discussions	Nearly all participants raised the importance of acknowledging that although they have dementia, there were many aspects of their lives that remained the same.	Not serious	High	High	High	High
Theme: PWD: Many participants described ways in which they covered up their dementia.							
1 (Pipon-Young 2012)	Interviews, group discussions	Reasons for this surrounded the uncertainty of others' reactions and perceptions of them. Participants described wishing others would keep seeing them as the person they always were and 'normal'.	Not serious	High	High	Low[3]	Low

Studies	Study design	Description	Methodological limitations	Relevance	Coherence	Adequacy	Confidence
Theme: PWD: Other people saw it as better to tell others that they had dementia, so they could understand their difficulties.							
1 (Pipon-Young 2012)	Interviews, group discussions	Other people saw it as better to tell others that they had dementia, so they could understand their difficulties.	Not serious	High	High	Low[3]	Low
Theme: PWD: Participants spoke of the importance of remaining independent, active and involved.							
1 (Pipon-Young 2012)	Interviews, group discussions	This could be achieved by finding a reason to keep fighting and not only focusing on deficits.	Not serious	High	High	Low[3]	Low
Theme: PWD: Many participants spoke of the importance of knowing other people with dementia and being able to share understandings through similar experiences.							
1 (Pipon-Young 2012)	Interviews, group discussions	Many participants spoke of the importance of knowing other people with dementia and being able to share understandings through similar experiences.	Not serious	High	High	Low[3]	Low
Theme: PWD: Participants described support from partners, friends, family, services, professionals, and through faith and spirituality.							
1 (Pipon-Young 2012)	Interviews, group discussions	Participants described support from partners, friends, family, services, professionals, and through faith and spirituality.	Not serious	High	High	Low[3]	Low
Theme: PWD: Resilience							
1 (Pipon-Young 2012)	Interviews, group discussions	There was a sense from participants that being diagnosed with dementia was not a helpless situation. There were still things they could do for themselves.	Not serious	High	High	Low[3]	Low
Theme: PWD: Participants discussed keeping their brains stimulated							
1 (Pipon-Young 2012)	Interviews, group discussions	Participants discussed keeping their brains stimulated.	Not serious	High	High	Low[3]	Low
Theme: PWD: Disconnection and isolation							
1 (Clemerson 2014)	Semi-structured interviews	A shared phenomenon of feeling isolated or disconnected from others emerged, which is heightened by a lack of age-appropriate services.	Serious[1]	High	High	Low[3]	Very low
Theme: PWD: Re-engaging in life following people's initial experience of disconnection and isolation.							

Studies	Study design	Description	Methodological limitations	Relevance	Coherence	Adequacy	Confidence
1 (Clemerson 2014)	Semi-structured interviews	Although disconnection was identified as a way of managing the sense of difference to others, it was recognised that this could not be sustained long term	Serious[1]	High	High	Low[3]	Very low
Theme: PWD: As people began to reconnect with others, their focus shifted.							
1 (Clemerson 2014)	Semi-structured interviews	Their focus shifted from concern with how they cope to concern with how their loved ones cope. Others focussed their attentions on contributing to the community and helping other people with dementia.	Serious[1]	High	High	Low[3]	Very low
Theme: PWD: The intention to regain control emerged as a common coping strategy in response to the experience of loss of agency.							
1 (Clemerson 2014)	Semi-structured interviews	The intention to regain control emerged as a common coping strategy in response to the experience of loss of agency.	Serious[1]	High	High	Low[3]	Very low
Theme: PWD: Dementia Service User Network (otherwise known as the 'Forget-Me-Nots') provide social comradeship and are a useful resource							
1 (Clayton-Turner 2015)	Interviews	Dementia Service User Network (otherwise known as the 'Forget-Me-Nots') provide social comradeship and are a useful resource	Serious[1]	High	High	Moderate[1]	Low
Theme: PWD: Making the most of life							
1 (Clayton-Turner 2015)	Interviews	Receiving a diagnosis of a life-limiting condition tends to concentrate the mind. It helps you recognise what is important, clarifying life goals and helping you identify things you want to do. Dementia forces you to make the most of every day, to live in the moment and cherish times of fun, intimacy and discovery. You find a new strength within and a depth to some relationships which become closer through the hard times.	Serious[1]	High	High	Moderate[1]	Low
Theme: PWD: Younger people living with dementia find YoungDementia UK very helpful.							
1 (Clayton-Turner 2015)	Interviews	Younger people living with dementia find YoungDementia UK very helpful.	Serious[1]	High	High	Moderate[1]	Low

Studies	Study design	Description	Methodological limitations	Relevance	Coherence	Adequacy	Confidence
Theme: Carer & PWD: Having dementia is frustrating, concerning and induces fear							
1 (Clayton-Turner 2015)	Interviews	Having dementia is frustrating, concerning and induces fear, and caring for a young person with dementia is stressful.	Serious[1]	High	High	Moderate[1]	Low
Theme: Carer: There is a lack of support for younger people living with dementia and their carers.							
1 (Clayton-Turner 2015)	Interviews	There is a lack of support for younger people living with dementia and their carers	Serious[1]	High	High	Moderate[1]	Low
Theme: Carer: When caring for a younger person living with dementia, key to coping and staying well is to carve out time for self							
1 (Clayton-Turner 2015)	Interviews	When caring for a younger person living with dementia, key to coping and staying well is to carve out time for self	Serious[1]	High	High	Moderate[1]	Low
Theme: Carer: Carers can receive support online at Talking Point, a peer support community run by Alzheimer's Society.							
1 (Clayton-Turner 2015)	Interviews	Carers can receive support online at Talking Point, a peer support community run by Alzheimer's Society	Serious[1]	High	High	Moderate[1]	Low
Theme: Carer: A diagnosis of dementia should be made before stopping work.							
1 (Clayton-Turner 2015)	Interviews	Otherwise, a person may not get their full pension. If a person stops working because of sickness, they may get their full pension. In addition, a diagnosis might enable the person to continue working at a reduced role or with support	Serious[1]	High	High	Moderate[1]	Low
Theme: Carer: Driving should be discussed.							
1 (Clayton-Turner 2015)	Interviews	Driving should be discussed	Serious[1]	High	High	Moderate[1]	Low

Studies	Study design	Description	Methodological limitations	Relevance	Coherence	Adequacy	Confidence
Theme: Carer: Becoming involved with research is advantageous for younger people living with dementia and their carers.							
1 (Clayton-Turner 2015)	Interviews	Becoming involved with research is advantageous for younger people living with dementia and their carers	Serious[1]	High	High	Moderate[1]	Low
Theme: Carer: Younger people living with dementia benefit from having relationships that are allowed to develop.							
1 (Clayton-Turner 2015)	Interviews	Younger people living with dementia benefit from having relationships that are allowed to develop	Serious[1]	High	High	Moderate[1]	Low

1. Theme only identified in studies at moderate risk of bias.
2. This is the only UK study that addresses this theme.
3. This is the only UK study that addresses this theme, and contains only a very small numbers of participants.

Themes identified for a walking group for younger people living with dementia and their carers

Studies	Study design	Description	Methodological limitations	Relevance	Coherence	Adequacy	Confidence
Theme: PWD: The walking group created supportive and positive relationships, bringing closeness, friendship and compassion.							
1 (Hegarty 2014)	focus group interview, questionnaire	The walking group created supportive and positive relationships, bringing closeness, friendship and compassion.	Not serious	High	High	Low[1]	Low
Theme: PWD: Group members were clear about the benefits to partners							
1 (Hegarty 2014)	focus group interview, questionnaire	Group members were clear about the benefits to partners.	Not serious	High	High	Low[1]	Low
Theme: PWD: Some talked about the disadvantages of having a large walking group.							

Studies	Study design	Description	Methodological limitations	Relevance	Coherence	Adequacy	Confidence
1 (Hegarty 2014)	focus group interview, questionnaire	Some talked about the disadvantages of having a large walking group.	Not serious	High	High	Low[1]	Low
Theme: Carer: Through the spouses' questionnaire, partners reported some positive impact on physical health and communication skills, and a substantial positive impact on mood.							
1 (Hegarty 2014)	focus group interview, questionnaire	Through the spouses' questionnaire, partners reported some positive impact on physical health and communication skills, and a substantial positive impact on mood.	Not serious	High	High	Low[1]	Low

1. This is the only UK study that addresses this theme, and contains only a very small numbers of participants.

Themes identified for a day service for younger people living with dementia

Studies	Study design	Description	Methodological limitations	Relevance	Coherence	Adequacy	Confidence
Theme: A sense of belonging							
1 (Higgins 2010)	Interviews	To feel part of a valued group, to maintain or form important relationships. An opportunity to simply 'be myself' and 'not pretend' are important to evaluative outcomes of a successful service.	Not serious	High	High	Low[1]	Low
Theme: ACE club provided a sense of achievement.							
1 (Higgins 2010)	Interviews	It enabled members to reach valued goals to the satisfaction of self and/or others. In considering this sense and its place in their life, ACE club members took a broad viewpoint on inclusion, which included a focus on physical rehabilitation to promote health and well-being, and supported practical strategies for daily living to promote confidence and reaffirm roles within the home.	Not serious	High	High	Low[1]	Low

Studies	Study design	Description	Methodological limitations	Relevance	Coherence	Adequacy	Confidence
Theme: ACE club enabled members to talk through their problems							
1 (Higgins 2010)	Interviews	ACE club enabled members to talk through their problems.	Not serious	High	High	Low[1]	Low
Theme: ACE club provides a sense of purpose							
1 (Higgins 2010)	Interviews	ACE club provides a sense of purpose.	Not serious	High	High	Low[1]	Low
Theme: A sense of security							
1 (Higgins 2010)	Interviews	To feel safe physically, psychologically, existentially. Many of the responses shared by members in the evaluation reinforce a sense of security on many levels. However, the inclusive nature of the membership of the ACE club strengthened the sense of security for the wider family and this was seen as a vital part of the service and the meaning that it held for members. The evaluation process demonstrated that group cohesion provided a sense of security for its membership where 'permission' to be vulnerable within a supportive environment was essential to human growth. Without this sense of security, some members feared that they would simply have to return to smaller family networks where their role and status may not be so well supported.	Not serious	High	High	Low[1]	Low
Theme: A sense of significance							
1 (Higgins 2010)	Interviews	To feel that you 'matter' and are accorded value and status. Interestingly, this was the 'sense' that was evaluated by the ACE club members as being the most important. Significance was experienced on a number of levels and with multiple meanings. The ACE club members valued the opportunities to speak at local, regional and national conferences with their campaigning voice for younger people with dementia, helping to spark and inform the development of a	Not serious	High	High	Low[1]	Low

Studies	Study design	Description	Methodological limitations	Relevance	Coherence	Adequacy	Confidence
		number of service philosophies and initiatives across the country, as well as inspire similar clubs in Australia, namely CALM and ConnexUS in Adelaide, South Australia. Additionally, members saw the significance of being involved in teaching clinical psychology students and student nurses. This sense of significance cascaded through their lives both at home and within the wider community and enhanced their experience of living and reaffirmed their sense of self.					
Theme: ACE club was felt to slow down the progression of dementia							
1 (Higgins 2010)	Interviews	ACE club was felt to slow down the progression of dementia.	Not serious	High	High	Low[1]	Low
1. This is the only UK study that addresses this theme, and contains only a very small numbers of participants.							

Themes identified for a lunchtime social group for younger women living with dementia ('Ladies who Lunch')

Studies	Study design	Description	Methodological limitations	Relevance	Coherence	Adequacy	Confidence
Theme: PWD: Ladies who Lunch provided value to those attending it							
1 (Johnson 2008)	Written and verbal feedback	Ladies who Lunch provided companionship, a relaxing atmosphere, was enjoyable and was valued by bot the women and their carers.	Serious[1]	High	High	Moderate[2]	Low
Theme: Carer: Ladies who Lunch gives younger women living with dementia greater confidence							
1 (Johnson 2008)	Written and verbal feedback	Ladies who Lunch gives younger women living with dementia greater confidence.	Serious[1]	High	High	Moderate[2]	Low
1. Written and verbal feedback is likely to result in data from motivated participants. Less motivated participants' views might not have been forthcoming and those views could be valuable. There was no before and during comparison. Inclusion and exclusion criteria are not provided; nor are characteristics of the participants. 2. This is the only UK study that addresses this theme.							

G.14 Assessing and managing comorbidities

G.14.1 Assessing and treating intercurrent illness in people living with dementia

- Are there effective methods for assessing intercurrent illness in people living with dementia that are different from those already in use for people who do not have dementia?
- Are there effective methods for treating intercurrent illness in people living with dementia that are different from those already in use for people who do not have dementia?

G.14.1.1 Assessing intercurrent illness

Observer rated versus self-report pain assessment

Pain Assessment in Advanced Dementia (PAINAD) and Numerical Rating Scale (NRS)

Quality assessment							No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Other considerations	Cognitive impairment (CI)	Cognitively intact (non CI)	Summary of results	
Outcome : Presence of pain as assessed by PAINAD and NRS										
Mosele (2012)	Prospective cohort	Serious[1]	Not serious	N/A	Not serious	None	310	290	PAINAD MD 0.70 (0.26, 1.14)	Moderate
Mosele (2012)	Prospective cohort	Serious[1]	Not serious	N/A	Serious[2]	None	310	290	NRS MD = 0.30 (-0.25 to 0.85)	Low
Prevalence of pain										
Mosele (2012)	Prospective cohort	Serious[1]	Not serious	N/A	Serious[3]	None	310	290	PAINAD	Low

Quality assessment							No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Other considerations	Cognitive impairment (CI)	Cognitively intact (non CI)	Summary of results	
									RR 1.39 (1.20, 1.62)	
Mosele (2012)	Prospective cohort	Serious[1]	Not serious	N/A	Serious[3]	None	310	290	NRS RR 1.19 (1.00, 1.41)	Low

1 *Risk of selection bias in study*
2 *Non-significant result*
3 *95% CI Crosses one line of a defined MID interval*

Observational versus self-report pain assessmentNon Communicative Patients Pain Assessment (NOPPAIN), Numerical Rating Scale (NRS) and Verbal Descriptor Scale (VDS)

Quality assessment							No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Other considerations	Cognitive impairment (CI)	Cognitively intact (non CI)	Summary of results	
Outcome : Presence of pain as assessed by NOPPAIN, NRS and VDS Relationship between observational (NOPPAIN) scores and self-report scores										
Correlation of NOPPAIN intensity with how much pain participants report										
Horgas (2012)	Cross sectional	Serious[1]	Not serious	N/A	Serious[2]	None	20	20	CI group VDS *r*=0.05, p= non sig NRS *r*=0.16, p=non sig	Low

Quality assessment							No of patients		Effect estimate	
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Other considerations	Cognitive impairment (CI)	Cognitively intact (non CI)	Summary of results	Quality
									Non CI group VDS *r*=0.66, p<0.001 NRS r=0.66, p<0.001	
Correlation of NOPPAIN intensity with total no of pain indicators observed										
Horgas (2012)	Cross sectional	Serious[1]	Not serious	N/A	Serious[2]	None	20	20	CI group r=0.63, p<0.001 Non CI group r=0.65, p<0.001	Low

[1]*Risk of selection bias*
[2]*Small sample size*

Observational versus self-report pain assessment

Pain Assessment in Advanced Dementia (PAINAD) and Numerical Rating Scale (NRS)

Quality assessment							No of patients		Effect estimate	
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Other considerations	Cognitive impairment (CI)	Cognitively intact (non CI)	Summary of results	Quality
Outcome : Correlation between PAINAD and NRS										

Quality assessment							No of patients		Effect estimate	
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Other considerations	Cognitive impairment (CI)	Cognitively intact (non CI)	Summary of results	Quality
De Waters (2008)	Correlational	Serious[1]	Serious[2]	N/A	Serious[3]	None	12	13	CI group r^a=0.735 p<0.001 Non CI group r=0.915 p<0.001	Very low

[1]*Risk of selection bias*
[2]*Sub sample drawn from larger populatin of elderly hip fracture patients*
[3]*Small sample size*
(a) Pearsons's correlation coefficient

Observational versus observational and self-report pain assessment

Rotterdam Elderley Pain Observation Scale, PAINAD and NRS (REPOS versus PAINAD and NRS)

Quality assessment							No of patients		Effect estimate	
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Other considerations	Cognitive impairment (CI)	Cognitively intact (non CI)	Summary of results	Quality
Outcome : Correlation between (REPOS versus PAINAD and NRS)										
Van Herk (2009)	Case control	Serious[1,2]	Not serious	N/A	Not serious	None	124	50	CI group PAINAD rs^a=0.75 (0.66 to 0.82) NRS-nurse rs =0.19 (0.01 to 0.35)	Low

Quality assessment							No of patients		Effect estimate	
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Other considerations	Cognitive impairment (CI)	Cognitively intact (non CI)	Summary of results	Quality
									Non CI group PAINAD rs=0.61 (0.40 to 0.76) NRS-nurse rs =0.36 (0.09 to 0.58)	
Comparison of pain scores: Median REPOS scores during painful activity										
Van Herk (2009)	Case control	Serious[1,2]	Serious[3]	N/A	Not serious	None	124	50	CI group= 5 (IQR 3 to 6) Non CI group =4 (IQR 3 to 5) (p=0.0002)[b]	Very low

[1] Risk of selection bias
[2] Selective reporting of methods
[3]Control group included people with MMSE≥18. Cannot be certain that this may have included people with Mild cognitive impairment
(a) Spearman's rank correlation coefficient
(b) Based on two-way ANOVA

Observational versus observational and observational pain assessment versus self-report (Abbey pain scale versus PAINAD and NOPPAIN versus self-report)

Quality assessment							No of patients		Effect estimate	
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Other considerations	Cognitive impairment (CI)	Cognitively intact (non CI)	Summary of results	Quality
Outcome : Correlation between observational ratings and self-report ratings of pain intensity										
Lukas (2013)	Retrospective cohort	Serious[1]	Not serious	N/A	Not serious	None	49	59	CI group Abbey r=0.563 (p<0.001) PAINAD r=0.532 (p<0.001) NOPPAIN r=0.680 (p<0.001) Non CI group Abbey r=0.314 (p=0.015) PAINAD r=0.241 (p=0.066) NOPPAIN r=0.320 (p=0.013)	Moderate
Agreement of self-reported and observational-rated pain										

Quality assessment							No of patients		Effect estimate	
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Other considerations	Cognitive impairment (CI)	Cognitively intact (non CI)	Summary of results	Quality
Lukas (2013)	Retrospective cohort	Serious[1]	Not serious	N/A	Not serious	None	49	59	CI group Abbey 78.3% PAINAD 73.3% NOPPAIN 80.0% Non CI group Abbey 66.1% PAINAD 66.1% NOPPAIN 69.2%	Moderate

[1]*Risk of selection bias*

Falls assessment versus functional assessment: Berg Balance Scale (BBS)

Quality assessment							No of patients		Effect estimate	
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Other considerations	Cognitive impairment (CI)	Cognitively intact (non CI)	Summary of results	Quality
Outcome : Performance on BBS										

Quality assessment							No of patients		Effect estimate	
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Other considerations	Cognitive impairment (CI)	Cognitively intact (non CI)	Summary of results	Quality
Kato-Narita (2011)	Case control	Serious[1]	Not serious	N/A	Serious[2]	None	48	40	Mean difference in scores CI group =51.3; Non CI group=53.1 (p=0.001) MD = -1.80 (-3.06 to -0.54)	Low
Correlation between number of falls recorded in last 12 months and scores on BBS										
Kato-Narita (2011)	Case control	Serious[1]	Not serious	M/A	Serious[2]	None	23[a]	40	CI group r= -0.613 (p=0.045) Non CI group r=0.383 (p=0.015)	Low

[1] Risk of selection bias level
[2]Based on small sample and sup population of wider sample
(a) Sample based on subpopulation classified as mild AD (classified by Clinical Dementia Rating (CDR)

Delirium assessment

Quality assessment							No of patients		Effect estimate	
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Other considerations	Cognitive impairment (CI)	Cognitively intact (non CI)	Summary of results	Quality
AUCa for DRS versus DSM-5										
Sepulveda (2015)	Cross-sectional	Serious[1]	Not serious	N/A	Serious[2]	None	85	40	CI group = 87.03%; Non CI group = 98.86% MD 11.83 (3.07 to 20.59)	Low
AUC for DRS versus ICD-10										
Sepulveda (2015)	Cross-sectional	Serious[1]	Not serious	N/A	Serious[2]	None	85	40	CI group = 86.69%; Non CI group = 97.37% MD 10.68 (1.62 to 19.74)	Low
AUC for DRS versus DSM-III-R										
Sepulveda (2015)	Cross-sectional	Serious[1]	Not serious	N/A	Serious[2]	None	85	40	CI group = 88.55%; Non CI group = 100%	Low

Quality assessment							No of patients		Effect estimate	
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Other considerations	Cognitive impairment (CI)	Cognitively intact (non CI)	Summary of results	Quality
									MD 11.45 (3.02 to 19.88)	
AUC for DRS versus DSM-IV										
Sepulveda (2015)	Cross-sectional	Serious[1]	Not serious	N/A	Serious[2]	None	85	40	CI group = 88.29%; Non CI group = 100% MD 11.71 (3.44 to 19.98)	Low

[1]*Observational design, downgrade 1 level*
[2]*Based on small sample and sup population of wider sample*
AUC= Area under the curve

G.14.1.2 Management of intercurrent illness

Pain Management

Quality assessment							No of patients		Effect estimate	
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Other considerations	Intervention	Control	Summary of results	Quality
Change in PRN medication quantification scores per unit of assessment time (PACSLAC vs activity log) – 3 months										

Quality assessment							No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Other considerations	Intervention	Control	Summary of results	
Fuchs-Lacelle (2008)	Cluster RCT	Serious[1]	Not serious	N/A	Not serious	None	89	84	MD 0.005 (p value = 0.00)	Low
Nursing stress scale: total score (PACSLAC vs activity log) – 3 months										
Fuchs-Lacelle (2008)	Cluster RCT	Serious[1]	Not serious	N/A	Not serious	None	89	84	MD -6.10 (p value = 0.04)	Low
Overall pain intensity: MOBID-2 (stepwise-treatment vs usual care) – 8 weeks										
Sandvik (2014)	Cluster RCT	Serious[2]	Not serious	N/A	Not serious	None	164	163	-1.393 (p value < 0.001)	Moderate
NPI-NH total score (stepwise-treatment vs usual care) – 8 weeks										
Husebo (2014)	Cluster RCT	Serious[2]	Not serious	N/A	Not serious	None	142	156	-9.6 (p value < 0.001)	Moderate

[1]*No blinding of intervention or assessment, high dropout rate*
[2]*No adequate description of usual care*

Delirium

Quality assessment							No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Other considerations	Intervention	Control	Summary of results	
Barthel Index (Intervention versus control) – 30 days										
Kolanowski (2011)	RCT	Serious[1]	Not serious	N/A	Serious[2]	None	11	5	MD 4.33 (p value (group/time	Very low

Quality assessment							No of patients		Effect estimate	
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Other considerations	Intervention	Control	Summary of results	Quality
									interaction) = 0.001)	
Confusion Assessment Method (Intervention versus control) – 30 days										
Kolanowski (2011)	RCT	Serious[1]	Not serious	N/A	Serious[2]	None	11	5	MD -0.17 (p value (group/time interaction) = 0.1128)	Very low
Delirium Rating Scale (Intervention versus control) – 30 days										
Kolanowski (2011)	RCT	Serious[1]	Not serious	N/A	Serious[2]	None	11	5	MD -1.80 (p value (group/time interaction) = 0.0842)	Very low
MMSE (Intervention versus control) – 30 days										
Kolanowski (2011)	RCT	Serious[1]	Not serious	N/A	Serious[2]	None	11	5	MD 0.59 (p value (group/time interaction) = 0.0298)	Very low

[1]*No blinding of intervention or assessment, lack of clarity in methods*
[2]*Sample size of only 16 people*

Hip fracture

Quality assessment							No of patients	Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Other considerations		Summary of results	
Barthel Index (Intervention versus control) – 30 days									
Stenvall (2007)	Cluster RCT	Not serious	Not serious	N/A	Serious[2]	None	199	Full population: IRR 0.38 (0.20, 0.76) Dementia sub-population: IRR 0.07 (0.01, 0.57)	Moderate
Mortality (Enhanced inpatient care vs conventional care) – 12 months									
1: Smith (2015)	SR of RCTs	Serious[1]	Not serious	N/A	Serious[2]	None	47	OR 2.25 (0.67, 7.61)	Low
Personal activities of daily living independence (Enhanced inpatient care vs conventional care) – 12 months									
1: Smith (2015)	SR of RCTs	Serious[1]	Not serious	N/A	Very serious[3]	None	47	OR 4.62 (0.18, 119.63)	Very low
Mortality (Enhanced inpatient and home care vs conventional care) – 12 months									
2: Smith (2015)	SR of RCTs	Serious[1]	Not serious	N/A	Very serious[3]	None	177	OR 1.07 (0.47, 2.45)	Very low
Activities of daily living (Enhanced inpatient and home care vs conventional care) – 12 months									
1: Smith (2015)	SR of RCTs	Serious[1]	Not serious	N/A	Not serious	None	36	MD 25.40 (10.89, 39.91)	Moderate
Incidence of falls (Enhanced inpatient and home care vs conventional care) – 12 months									
1: Smith (2015)	SR of RCTs	Serious[1]	Not serious	N/A	Very serious[3]	None	36	OR 0.20 (0.01, 4.47)	Very low
Cumulative incidence of delirium (Geriatrician-led inpatient management vs orthopaedic-led inpatient management) – acute hospitalisation									
1: Smith (2015)	SR of RCTs	Serious[1]	Not serious	N/A	Very serious[3]	None	126	OR 0.73 (0.22, 2.38)	Very low

[1]Lack of reporting of trial methods
[2]Non-significant result
[3]95% CI crosses two lines of a defined MID interval

Falls

Quality assessment							No of patients		Effect estimate	
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Other considerations	Intervention	Control	Summary of results	Quality
Community: Home-based exercise versus usual care – mean number of falls										
2 (Pitkälä, Wesson)	RCT	Serious	Not serious	Not serious	Not serious	None	74	74	MD -1.07 (-1.78, -0.36)	Moderate
Community: Home-based exercise versus usual care – proportion of people falling										
2 (Pitkälä, Wesson)	RCT	Serious	Not serious	Not serious	Serious[2]	None	74	74	RR 0.69 (0.51, 0.93)	Low
Community: Home-based exercise versus usual care – Zarit Burden Score										
2 (Suttanon, Wesson)	RCT	Serious	Not serious	Not serious	Serious[3]	None	26	32	MD 4.02 (-3.16, 11.19)	Low
Community: Group-based exercise versus usual care – mean number of falls										
Pitkälä (2013)	RCT	Not serious	Not serious	N/A	Serious[3]	None	60	63	MD -1.03 (-2.19, 0.13)	Moderate
Community: Group-based exercise versus usual care – proportion of people falling										
Pitkälä (2013)	RCT	Not serious	Not serious	N/A	Serious[2]	None	60	63	RR 0.68 (0.50, 0.94)	Moderate
Exercise versus usual care – proportion of people falling										
7: Chan (2015)	SR of RCTs	Not serious	Not serious	Serious	Serious[2]	Some contacted authors did not return study data	372	316	RR 0.68 (0.51, 0.91)	Moderate
Exercise versus usual care – proportion of people with fractures										

Quality assessment							No of patients		Effect estimate	
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Other considerations	Intervention	Control	Summary of results	Quality
2: Chan (2015)	SR of RCTs	Serious	Not serious	Not serious	Very serious[4]	Some contacted authors did not return study data	185	119	RR 1.47 (0.56, 3.81)	Very low
Meta-regression for effect of prevalence of dementia on effect size of interventions										
43: Oliver (2006)	SR	Serious	Not serious	Serious	Serious[3]	None	Not reported		p value - rate ratio for falls: 0.72 p value – relative risk for fallers: 0.87 p value - rate ratio for fractures: 0.18	Very low
Multifactorial intervention versus usual care – proportion of people falling										
Shaw (2003)	RCT	Not serious	Serious[1]	N/A	Not serious	None	130	144	RR 0.92 (0.81, 1.05)	Moderate
Multifactorial intervention versus usual care – fractured neck of femur										
Shaw (2003)	RCT	Not serious	Serious[1]	N/A	Very serious[4]	None	130	144	RR 0.55 (0.21, 1.43)	Very low
Multifactorial intervention versus usual care – fall-related A&E attendance										
Shaw (2003)	RCT	Not serious	Serious[1]	N/A	Serious[2]	None	130	144	RR 1.25 (0.91, 1.72)	Low
Multifactorial intervention versus usual care – fall-related hospital admission										
Shaw (2003)	RCT	Not serious	Serious[1]	N/A	Very serious[4]	None	130	144	RR 1.11 (0.61, 2.00)	Very low
Multifactorial intervention versus usual care – mortality										
Shaw (2003)	RCT	Not serious	Serious[1]	N/A	Very serious[4]	None	130	144	RR 1.03 (0.65, 1.64)	Very low
Home-based technology intervention – proportion of people falling										

Quality assessment							No of patients		Effect estimate	
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Other considerations	Intervention	Control	Summary of results	Quality
Tchalla (2013)	RCT	Not serious	Not serious	N/A	Serious[2]	None	49	47	OR 0.37 (0.15, 0.88)	Moderate

[1]Contains patients with cognitive impairment but no diagnosis of dementia
[2]95% CI crosses one line of a defined MID interval
[3]Non-significant result
[4]95% CI crosses one line of a defined MID interval

G.14.2 Management strategies for people living with dementia and co-existing physical long term conditions

- What are the optimal management strategies (including treatments) for people living with dementia with co-existing physical long term conditions?

G.14.2.1 Hypertension

Quality assessment							No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Other considerations	Telmisartan (n=10)	Amlodipine (n=10)	Summary of results	
Clinical progression of comorbidity & associated symptoms										
Mean difference in systolic BP at 6 months (PPAR versus CCB)										
Kume (2012)	Randomised open label trial	Serious[1]	Not serious	N/A	Very serious[2]	None	10	10	MD 2.00 (-7.64, 11.64)	Very low
Mean difference in diastolic BP at 6 months (PPAR versus CCB)										
Kume (2012)	Randomised open label trial	Serious[1]	Not serious	N/A	Very serious[2]	None	10	10	MD -2.00 (-8.20, 4.20)	Very low
Mean difference in pulse rate at 6 months (PPAR versus CCB)										
Kume (2012)	Randomised open label trial	Serious[1]	Not serious	N/A	Very serious[2]	None	10	10	MD 2.00 (-1.61, 5.61)	Very low
Clinical outcomes, including cognitive, functional, behavioural ability										
Mean difference in MMSE at 6 months (PPAR versus CCB)										
Kume (2012)	Randomised open label trial	Serious[1]	Not serious	N/A	Very serious[2]	None	10	10	MD 0.00 (-3.10, 3.10)	Very low
Mean difference in ADAS-Cog at 6 months (PPAR versus CCB)										
Kume (2012)	Randomised open label trial	Serious[1]	Not serious	N/A	Very serious[2]	None	10	10	MD -1.10 (-6.32, 4.12)	Very low
Mean difference in WMS-R (logical- memory) at 6 months (PPAR versus CCB)										
Kume (2012)	Randomised open label trial	Serious[1]	Not serious	N/A	Very serious[2]	None	10	10	MD 3.00 (-0.18, 6.18)	Very low

1. Downgrade 1 level selective reporting of methods

Quality assessment							No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Other considerations	Telmisartan (n=10)	Amlodipine (n=10)	Summary of results	

2. Downgrade 2 levels; small sample size and wide confidence intervals

Quality assessment							No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Other considerations	Relative-HBPM (n=60)	ABPM (n=60)	Summary of results	
Clinical progression of comorbidity & associated symptoms										
Mean difference in systolic BP after 3 days (R-HBPM versus 24-h ABPM)										
Plichart (2013)	Randomised open comparative cross over study	Serious[1]	Not serious	N/A	Serious[2]	None	60	60	MD 11.30 (4.61, 17.99)	Low
Mean difference in diastolic BP after 3 days (R-HBPM versus 24-h ABPM)										
Plichart (2013)	Randomised open comparative cross over study	Serious[1]	Not serious	N/A	Serious[2]	None	60	60	MD 1.00 (-2.76, 4.76)	Low
Mean difference in systolic BP after 3 days (R-HBPM versus day ABPM)										
Plichart (2013)	Randomised open comparative cross over study	Serious[1]	Not serious	N/A	Serious[2]	None	60	60	MD 9.70 (3.08, 16.32)	Low
Mean difference in diastolic BP after 3 days (R-HBPM versus day ABPM)										
Plichart (2013)	Randomised open comparative cross over study	Serious[1]	Not serious	N/A	Serious[2]	None	60	60	MD 0.00 (-3.76, 3.76)	Low

1. Downgrade 1 level, crossover comparative design
2. Short follow up period, 3 days

G.14.2.2 Cardiovascular disease

Quality assessment							No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Other considerations	Vascular care (n=50)	Standard care (n=44)	Summary of results	
Clinical progression of comorbidity & associated symptoms										
Mean difference in change over 2 years systolic BP (SC versus VC)										
Richard 2012	Randomised controlled trial	Not serious	Not serious	N/A	Serious[1]	None	50	44	MD -4.12 (-14.75, 6.16)	Moderate
Mean difference in change over 2 years diastolic BP (SC versus VC)										
Richard 2012	Randomised controlled trial	Not serious	Not serious	N/A	Serious[1]	None	50	44	MD -1.97 (-8.21, 4.26)	Moderate
Mean difference in change over 2 years HBA1C (SC versus VC)										
Richard 2012	Randomised controlled trial	Not serious	Not serious	N/A	Serious[1]	None	50	44	MD 0.20 (-0.08, 0.48)	Moderate
Mean difference in change over 2 years total cholesterol (SC versus VC)										
Richard 2012	Randomised controlled trial	Not serious	Not serious	N/A	Serious[1]	None	50	44	MD -0.94 (-1.43, -0.45)	High
Mean difference in change over 2 years HDL cholesterol (SC versus VC)										
Richard 2012	Randomised controlled trial	Not serious	Not serious	N/A	Serious[1]	None	50	44	MD -0.02 (-0.17, 0.13)	Moderate
Mean difference in change over 2 years LDL cholesterol over 2 years (SC versus VC)										
Richard 2012	Randomised controlled trial	Not serious	Not serious	N/A	Serious[1]	None	50	44	MD -0.90 (-1.44, -0.36)	High
Clinical outcomes, including cognitive, functional, behavioural ability										
Mean difference in change over 2 years MMSE (SC versus VC)										
Richard 2012	Randomised controlled trial	Not serious	Not serious	N/A	Serious[1]	None	50	44	MD -0.55 (-3.12, 2.02)	Moderate

Quality assessment							No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Other considerations	Vascular care (n=50)	Standard care (n=44)	Summary of results	
Mean difference in change over 2 years IDDAD (SC versus VC)										
Richard 2012	Randomised controlled trial	Not serious	Not serious	N/A	Serious[1]	None	50	44	MD 2.71 (-3.14, 8.56)	Moderate
Mean difference in change over 2 years Revised MBPC (SC versus VC)										
Richard 2012	Randomised controlled trial	Not serious	Not serious	N/A	Serious[1]	None	50	44	MD 4.54 (-1.39, 10.49)	Moderate

1. Non-significant result

G.14.2.3 Diabetes

Quality assessment							No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Other considerations	Pioglitazone (n=21)	No drug (n=21)	Summary of results	
Clinical progression of comorbidity & associated symptoms										
Mean difference in fasting plasma glucose at 6 months (Pioglitazone versus Control)										
Sato 2011	Randomised open controlled trial	Not serious	Not serious	N/A	Very serious[1]	None	21	21	MD -0.50 (-1.14, 0.14)	Low
Mean difference in HBA1c at 6 months (Pioglitazone versus Control)										
Sato 2011	Randomised open controlled trial	Not serious	Not serious	N/A	Very serious[1]	None	21	21	MD -0.10 (-0.68, 0.48)	Low
Mean difference in fasting insulin at 6 months (Pioglitazone versus Control)										
Sato 2011	Randomised open controlled trial	Not serious	Not serious	N/A	Very serious[1]	None	21	21	MD -0.80 (-2.32, 0.72)	Low

Quality assessment							No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Other considerations	Pioglitazone (n=21)	No drug (n=21)	Summary of results	
Clinical outcomes, including cognitive, functional, behavioural ability										
Mean difference in MMSE at 6 months (Pioglitazone versus Control)										
Sato 2011	Randomised open controlled trial	Not serious	Not serious	N/A	Very serious[1]	None	21	21	MD-1.50 (-0.67, 3.67)	Low
Mean difference in ADAS-Cog at 6 months (Pioglitazone versus Control)										
Richard 2012	Randomised controlled trial	Not serious	Not serious	N/A	Very serious[1]	None	21	21	MD -3.30 (-6.86, 0.26)	Low
Mean difference in WMS-R logical memory at 6 months (Pioglitazone versus Control)										
Richard 2012	Randomised controlled trial	Not serious	Not serious	N/A	Very serious[1]	None	21	21	MD 2.40 (-0.13, 4.93)	Low

1. Downgrade 2 levels, non-significant effect and small sample size

G.14.2.4 Incontinence

Quality assessment							No of patients (n=74)		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Other considerations	IST programme (n=44)	Control group (n=30)	Summary of results	
Clinical progression of comorbidity & associated symptoms										
No of participants showing decreased incontinence at 6 months (IST versus control)										
Jirovec (2001)	RCT	Serious[1]	Not serious	N/A	Serious[2]	None	28/44	15/30	RR 1.27 (0.83, 1.94)	Low
Mean incontinence frequency at 6 months (IST versus control)										
Jirovec (2001)	RCT	Serious[1]	Not serious	N/A	Serious[3]	None	44	30	MD -0.12 (-0.27, 0.03)	Low
Clinical outcomes, including cognitive, functional, behavioural ability										
Mean difference in mental status (based on SPMSQ) score at 6 months IST versus control (IST versus control)										

Quality assessment							No of patients (n=74)		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Other considerations	IST programme (n=44)	Control group (n=30)	Summary of results	
Jirovec (2001)	RCT	Serious[1]	Not serious	N/A	Serious[3]	None	44	30	MD -0.46 (-1.48, 0.56)	Low
Mean difference in composite mobility score at 6 months (IST versus control)										
Jirovec (2001)	RCT	Serious[1]	Not serious	N/A	Serious[3]	None	44	30	MD 0.94 (-0.90, 2.78)	Low

1. Poorly reported study with unclear methods
2. 95% CI crosses one line of a defined MID interval
3. Non-significant result

Quality assessment							No of patients (N=19)		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Other considerations	Prompted voiding (n=9)	Control group (n=10)	Summary of results	
Clinical progression of comorbidity & associated symptoms										
Mean %ge reduction in all incontinent episodes per day (PV versus control) at 8 weeks										
Engberg (2002)	RCT	Serious[1]	Not serious	N/A	Serious[2]	None	9	10	MD 19.8 (-10.49 to 50.09)	Low
Mean %ge reduction in daytime incontinent episodes per day (PV versus control) at 8 weeks										
Engberg (2002)	RCT	Serious[1]	Not serious	N/A	Serious[2]	None	9	10	MD 12.8 (-21.55 to 47.15)	Low
Mean %ge reduction in daytime wet (PV versus control) at 8 weeks										
Engberg (2002)	RCT	Serious[1]	Not serious	N/A	Serious[2]	None	9	10	MD 8.5 (-28.35 to 45.35)	Low
Mean %ge reduction in day & night time wet (PV versus control) at 8 weeks										

Quality assessment							No of patients (N=19)		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Other considerations	Prompted voiding (n=9)	Control group (n=10)	Summary of results	
Engberg (2002)	RCT	Serious[1]	Not serious	N/A	Serious[2]	None	9	10	MD 17.60 (-14.58 to 49.78)	Low
Mean number of self-initiated toilets per day (PV versus control) at 8 weeks										
1.	RCT	Serious[1]	Not serious	N/A	Serious[2]	None	9	10	MD 1.20 (-2.20 to 4.60)	Low

1. Crossover aspect, participants in control crossed over to complete experimental phase
2. Small sample size with non-significant result

Quality assessment							No of patients		Effect estimate	Quality
No of studies	Design	Risk of bias	Indirectness	Inconsistency	Imprecision	Other considerations	Timed voiding (n=102	Control (n=89	Summary of results	
Clinical progression of comorbidity & associated symptoms										
Reduction in incidence of daytime incontinence after 2 months (TV versus usual care)										
Ostaskiewicz (2010)	Systematic review	Serious[1]	Not serious	N/A	Serious[2]	None	40/102	26/89	RR 1.34 (0.90 to 2.01)	Low
Reduction in incidence of night time incontinence after 2 months (TV versus usual care)										
Ostaskiewicz (2010)	Systematic review	Serious[1]	Not serious	N/A	Serious[2]	None	39/95	18/79	RR 1.80 (1.12 to 2.89)	Moderate
Reduction in volume of incontinence (based on pad volume) after 2 months (TV versus usual care)										
Ostaskiewicz (2010)	Systematic review	Serious[1]	Not serious	N/A	Very serious[3]	None	16/65	11/45	RR 1.01 (0.52 to 1.96)	Very low

High quality systematic review, included one low quality RCT
1. Downgrade 1 level; inadequate reporting of methods of allocation
2. 95% CI crosses one line of a defined MID interval
3. 95% CI crosses two lines of a defined MID interval

G.14.2.5 Age-related hearing impairment

Number of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
ADL: ADCS-ADL (follow up 6 months – higher numbers favour intervention								
1 (Adrait 2017)	RCT	Serious[1]	N/A	Not serious	Serious[2]	36	MD 0.20 (-1.21, 1.61)	Low
ADL: ADCS-ADL (follow-up 12 months) – higher numbers favour intervention								
1 (Adrait 2017)	RCT	Serious[1]	N/A	Not serious	Serious[2]	36	MD 0.30 (-1.19, 1.79)	Low
Behavioural and psychological symptoms: NPI (follow up 6 months) – lower numbers favour intervention								
1 (Adrait 2017)	RCT	Serious[1]	N/A	Not serious	Serious[2]	36	MD -2.50 (-14.95, 9.95)	Low
Behavioural and psychological symptoms: NPI (follow-up 12 months) – lower numbers favour intervention								
1 (Adrait 2017)	RCT	Serious[1]	N/A	Not serious	Serious[2]	36	MD -14.30 (-30.95, 2.35)	Low
Carer burden: ZBI (follow up 6 months – lower numbers favour intervention								
1 (Adrait 2017)	RCT	Serious[1]	N/A	Not serious	Serious[2]	36	MD -3.90 (-14.32, 6.52)	Low
Carer burden: ZBI (follow-up 12 months) – lower numbers favour intervention								
1 (Adrait 2017)	RCT	Serious[1]	N/A	Not serious	Serious[2]	36	MD -5.40 (-14.48, 3.68)	Low
Quality of life: ADRQL (follow-up 6 months) – higher numbers favour intervention								
1 (Adrait 2017)	RCT	Serious[1]	N/A	Not serious	Serious[2]	32	MD 5.60 (-40.39, 51.59)	Low
Quality of life: ADRQL (follow up 12 months) – higher numbers favour intervention								
1 (Adrait 2017)	RCT	Serious[1]	N/A	Not serious	Not serious	32	MD 43.20 (0.68, 85.72)	Moderate

1. Partial crossover design
2. Non-signficant result

G.15 Managing mental health conditions alongside dementia

- RQ20: What are the optimal management strategies (including treatments) for people with dementia and an enduring mental health condition?

No GRADE or CERQual tables were produced for this review question

G.16 Palliative care

G.16.1 Palliative care

- What models of palliative care are effective for people with dementia

G.16.1.1 Qualitative evidence

Carer identified issues

Studies	Study design	Description	Methodological limitations	Relevance	Coherence	Adequacy	Confidence
Bereaved carer – meeting physical care needs							
Lawrence (2011)	Structured interviews	Ensuring adequate food and fluid intake, hygiene, toileting, dressing.	Serious[1]	High	High	High	Moderate
Bereaved carer – going beyond task-focused care							
Crowther (2013), Lawrence (2011), Moore 2017	Structured interviews, Unstructured interviews	End-of-life care was evaluated positively if it was felt that the professionals cared about their dying relative.	Serious[1]	High	High	High	Moderate
Crowther (2013), Treloar (2009)	Unstructured interviews, Mixed methodology	Getting to know individual's interests, sensitivities and preferences (including food preferences).	Serious[1]	High	High	High	Moderate
Bereaved carer –planning							
Dening (2012), Lawrence (2011)	Structured interviews	Advance directives and advance statements.	Serious[1]	High	High	High	Moderate
Lawrence (2011)	Structured interviews	Discussing treatment planning with families and the wider care team.	Serious[1]	High	High	High	Moderate

Studies	Study design	Description	Methodological limitations	Relevance	Coherence	Adequacy	Confidence
Lawrence (2011)	Structured interviews	Enabling family members to be present at the time of death.	Serious[1]	High	High	High	Moderate
Dening (2012)	Semi-structured interviews, focus groups	Family carers described how little happened routinely; they had to initiate and then "push" for services to be provided, these were unpredictable and fragmented	Serious[1]	High	High	High	Moderate
Bereaved carer – impact of hospitalisation							
Dening (2012), Treloar (2009)	Semi-structured interviews, focus groups	Not liking the hospital environment.	Serious[1]	High	High	High	Moderate
Crowther (2013)	Unstructured interviews	Dying on an open ward rather than finding a side room in a hospital.	Serious[1]	High	High	High	Moderate
Dening (2012)	Semi-structured interviews, focus groups	Carers described how acute hospital staff struggled to provide basic care. Carers perceived a lack of understanding, little compassion and low staffing levels	Serious[1]	High	High	High	Moderate
Bereaved carer - Knowing the person well and having a sense of their personal and social identity was said to enable carers and health-care professionals to make better informed best interests decisions on behalf of a person with dementia							
1 Lamahewa (2017)	Focus groups and semi-structured interviews	This was thought to be particularly pertinent at the end of life, when the person with dementia may not always able to verbally express themselves.	Not serious	High	High	High	High
Bereaved carer – Knowledge of dementia provides insight for decision making							
1 Lamahewa (2017)	Focus groups and semi-	A sense of preparedness, understanding and insight into the impact of dementia on the end of life seemed likely to have resulted in a greater level of acceptance amongst some carers, which was said to have a	Not serious	High	High	High	High

Studies	Study design	Description	Methodological limitations	Relevance	Coherence	Adequacy	Confidence
	structured interviews	powerful influence on decision making between families and practitioners.					
Current carer - Lack of familiarity of the person with dementia by health-care providers inadvertently leads to disease labelling							
1 Lamahewa (2017)	Focus groups and semi-structured interviews	Lack of familiarity of the person with dementia by health-care providers inadvertently leads to disease labelling, whereby the individuality and identity of the person is lost and they are defined by their disease. This was considered to be particularly relevant when a person with dementia is admitted to hospital where staff have no information about them.	Not serious	High	High	High	High
Current carer - When healthcare professionals do not communicate with carers because of poor communication or lack of time to involve the family, this can complicate decision making							
1 Lamahewa (2017)	Focus groups and semi-structured interviews	When healthcare professionals do not communicate with carers because of poor communication or lack of time to involve the family, this can complicate decision making	Not serious	High	High	High	High
Current carer - Family carers reported often having to retell the same narrative to different health-care professionals							
1 Lamahewa (2017)	Focus groups and semi-structured interviews	There was a sense of frustration due to the lack of continuity in some settings, even within the same care setting	Not serious	High	High	High	High
Current carer – Carers sometimes have doubts making decisions, particularly if there was not an up-to-date living will							
1 Lamahewa (2017)	Focus groups and semi-structured interviews	Often decisions were based on the family member's insight about/or knowledge of the values or preferences of the person with dementia. However, they expressed feelings of uncertainty in how to best meet the needs of their relative. Further complications resulted if formal discussion had not taken place or if legal arrangements were not in place	Not serious	High	High	High	High
Carer - Carers often held strong views regarding the perceived quality of care							

Studies	Study design	Description	Methodological limitations	Relevance	Coherence	Adequacy	Confidence
1 Moore (2017)	Interviews	Carers often held strong views regarding the perceived quality of care	Not serious	High	High	High	High
Carer - Carers valued continuity and receiving regular feedback about their relative's health condition and the progression of dementia							
1 Moore (2017)	Interviews	Carers valued continuity and receiving regular feedback about their relative's health condition and the progression of dementia	Not serious	High	High	Moderate[1]	Moderate
Carer – Planning - Being able to monitor services was important and reflected poor levels of trust in service providers							
2 Moore (2017) Dening (2012)	Interviews	The standards of social service staff would drop if they felt they were not being monitored by the family. (Family carers described how little happened routinely; they had to initiate and then "push" for services to be provided, these were unpredictable and fragmented)	Not serious	High	High	High	High
Carer – Carers were rarely informed about the dementia from diagnosis onwards through to the palliative stages							
1 Moore (2017)	Interviews	Carers' capacity to understand the progression of dementia and be involved and informed during advanced dementia relied on information provision throughout the different stages of dementia. At diagnosis, carers were rarely informed about the likely progression of dementia	Not serious	High	High	Moderate[1]	Moderate
Carer - The unpredictable course of dementia made it very challenging for carers to prepare for the end of life							
1 Moore (2017)	Interviews	Some were unsure about the value of early information about advanced stages of disease given the potentially unnecessary anxiety this might create	Not serious	High	High	Moderate[1]	Moderate
Carer – Carers valued timely and sensitive information provided by a knowledgeable professional and that was reinforced in writing							
1 Moore (2017)	Interviews	Some felt that the lack of basic information left them struggling to adapt to changes and feeling ill-prepared for symptoms that they later discovered were common in advanced dementia	Not serious	High	High	Moderate[1]	Moderate
Carer – End of life (EOL) plans were not started early enough							
1 Moore (2017)	Interviews	End of life plans were rarely initiated during the early stages of dementia preventing the person with	Not serious	High	High	Moderate[1]	Moderate

Studies	Study design	Description	Methodological limitations	Relevance	Coherence	Adequacy	Confidence
		dementia being involved in decision making. Sometimes the person with dementia was never informed of their diagnosis. EOL planning often occurred after admission to a care home or after a critical health event usually involving hospitalisation in the advanced stages of dementia. Carers often appreciated these conversations as they could be involved in care and feel that they had contributed to a plan to promote comfort care at EOL.					
Carer – Some carers were satisfied with EOL care if they felt adequately informed and involved, even when EOL care was not in accordance with advance care plans							
1 Moore (2017)	Interviews	Some carers were satisfied with EOL care if they felt adequately informed and involved, even when EOL care was not in accordance with advance care plans	Not serious	High	High	Moderate[1]	Moderate
Carer – Enabling family members to be present at the time of death							
2 Moore (2017), Lawrence (2011)	Interviews	For most, but not all, being present at EOL was important and some described vigils from hours to weeks, being with the person before they died.	Not serious	High	High	High	High
Carer – Carers often grieve for their relative before the person dies							
1 Moore (2017)	Interviews	Carers described grief as a staged process pre and post death with losses associated with dementia before death.	Not serious	High	High	Moderate[1]	Moderate
Carer – There was evidence of links between satisfaction with EOL care, the carer's capacity to influence the care being provided, and emotional consequences							
1 Moore (2017)	Interviews	Two carers who had not moved their relative from what they perceived as a poor quality care home, reported the lowest satisfaction. This was influenced by their guilt at not having done more to improve EOL care.	Not serious	High	High	Moderate[1]	Moderate
Carer – Participants discussed the failure of services to acknowledge their grief or to provide information about obtaining support							
1 Moore (2017)	Interviews	This was both prior to and after their relative's death.	Not serious	High	High	Moderate[1]	Moderate

Studies	Study design	Description	Methodological limitations	Relevance	Coherence	Adequacy	Confidence
Carer - Despite high levels of grief, many carers felt they did not need formal support or counselling and did not seek it.							
1 Moore (2017)	Interviews	Instead they described the benefits of their social network including friends, family or faith community. Some carers could not face their grief or the fact that their relative had dementia.	Not serious	High	High	Moderate[1]	Moderate
Carer – Carers who felt well informed about how dementia progressed, were regularly updated on their relative's health condition and felt involved appeared more satisfied with EOL care.							
1 Moore (2017)	Interviews	Those who failed to influence care that they perceived as poor reported high levels of grief after death and experienced guilt and regret. Admission to a care home was often associated with a loss of control and a need for heightened vigilance	Not serious	High	High	Moderate[1]	Moderate

Professional identified issues

Studies	Study design	Description	Methodological limitations	Relevance	Coherence	Adequacy	Confidence
Professional – meeting physical care needs							
Lawrence (2011)	Structured interviews	Identifying and responding to the physical care needs of the person with dementia.	Serious[1]	High	High	High	Moderate
Lawrence (2011)	Structured interviews	Pain control.	Serious[1]	High	High	High	Moderate
Lawrence (2011)	Structured interviews	Palliative care nurses were considered skilled in identifying and managing pain in patients with complex needs and were also sensitive to nausea and hallucinations in people with dementia at the end of life.	Serious[1]	High	High	High	Moderate
Professional – complex pathways of care							
Dening (2012)	Semi-structured interviews,	People with advanced dementia had complex medical and social needs requiring input from a number of agencies, but the coordination was poor	Serious[1]	High	High	High	Moderate

Studies	Study design	Description	Methodological limitations	Relevance	Coherence	Adequacy	Confidence
	focus groups						
Dening (2012)	Semi-structured interviews, focus groups	Out of hours staff often felt unsupported and lacking in access to key information	Serious[1]	High	High	High	Moderate
Professional – going beyond task-focused care							
Lawrence (2011)	Structured interviews	Risk of becoming entirely task-focused with little empathy.	Serious[1]	High	High	High	Moderate
Lawrence (2011),	Structured interviews	Getting to know individual's interests, sensitivities and preferences.	Serious[1]	High	High	High	Moderate
Professional – planning							
Lawrence (2011), Grisaffi (2010)	Structured interviews, Semi-structured interviews	People with dementia should be given the opportunity to plan for the future.	Serious[1]	High	High	High	Moderate
Lawrence (2011)	Structured interviews	Whether individuals should be transferred to hospital during the final stages of their life. Hospitalisation was a frequent occurrence despite agreement among care professionals that this was often inappropriate.	Serious[1]	High	High	High	Moderate
Lawrence (2011)	Structured interviews	Palliative care staff noted that professionals across care settings could be reluctant to withdraw active treatment in the absence of explicit planning or a clear consensus among the care team.	Serious[1]	High	High	High	Moderate
Grisaffi (2010)	Semi-structured interviews	Discontinuity of care.	Serious[1]	High	High	Moderate[2]	Low
Professional – Flexibility							

Studies	Study design	Description	Methodological limitations	Relevance	Coherence	Adequacy	Confidence
Davies (2014)	Semi-structured interviews	The growing number of guidelines, standards, rules and regulations placed upon professionals in health and social care makes palliative care standardised leaving no room for flexibility.	Serious[1]	High	High	High	Moderate
Grisaffi (2010)	Semi-structured interviews	GP's prior knowledge of the person with dementia is important in informing decisions. To help the person overcome the communication and capacity issues, relatives and carers are seen as an expert source of information regarding the person's wishes.	Serious[1]	High	High	Moderate[2]	Low
Davies (2014)	Semi-structured interviews	NHS Primary Care Trusts have no duty of care for people who are self-funding their care home.	Serious[1]	High	High	High	Moderate
Professional - systemisation							
Davies (2014), Grisaffi (2010)	Semi-structured interviews	Some routines are useful, such as certain meetings, pain assessment, when to stop pursuing certain treatments.	Serious[1]	High	High	High	Moderate
Professional – staff training to reduce the need to call for specialist help.							
Davies (2014)	Semi-structured interviews	Syringe driver training, checks when prescribing.	Serious[1]	High	High	High	Moderate
Dening (2012)	Semi-structured interviews, focus groups	Many, particularly hospice, ambulance staff and district nurses acknowledged they had received little or no training in dementia, in particular concerning communication and managing behavioural problems	Serious[1]	High	High	High	Moderate
Professional - in some cases, the lack of palliative care skills is not seen as a gap to be filled by the generalist, rather the responsibility of a specialist service							
Davies (2014)	Semi-structured interviews	Some district nurses and GPs feel that palliative care should be left to specialists.	Serious[1]	High	High	High	Moderate
Professional – lack of trust, fear of litigation, fear of blame and threats to speciality							

Studies	Study design	Description	Methodological limitations	Relevance	Coherence	Adequacy	Confidence
Davies (2014)	Semi-structured interviews	Managing both real and perceived risks can be a difficult challenge	Serious[1]	High	High	High	Moderate
Professional - difficulty in deciding when to start end-of-life care							
Grisaffi (2010)	Semi-structured interviews	The typically slow erratic decline and the indicators for starting the pathway could lead to either a person being on it for a long time or 'yo-yoing' on and off as their state fluctuated.	Serious[1]	High	High	Moderate[2]	Low

1. Theme only identified in studies at moderate or high risk of bias
2. Insufficient data to develop a full understanding of the phenomenon of interest

G.16.1.2 Quantitative evidence

Specialist palliative care team versus usual care

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
Palliative care plan developed							
1 (Ahronheim 2000)	Serious[1]	N/A	Not serious	Not serious	99	RR 5.84 (1.37, 25.02)	Moderate
Palliative care plan during hospitalisation							
1 (Ahronheim 2000)	Serious[1]	N/A	Not serious	Serious[2]	99	RR 5.31 (0.26, 107.77)	Low
Palliative care plan on discharge							
1 (Ahronheim 2000)	Serious[1]	N/A	Not serious	Not serious	96	RR 4.50 (1.03, 19.75)	Moderate
Decision to forgo enteral feeds							
1 (Ahronheim 2000)	Serious[1]	N/A	Not serious	Serious[2]	99	RR 0.80 (0.19, 3.38)	Low
Decision to forgo mechanical ventilation							
1 (Ahronheim 2000)	Serious[1]	N/A	Not serious	Serious[2]	99	RR 7.43 (0.39, 140.15)	Low
Decision to forgo intravenous lines							
1 (Ahronheim 2000)	Serious[1]	N/A	Not serious	Serious[2]	99	RR 5.31 (0.64, 43.84)	Low
Decision to forgo blood draws							

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
1 (Ahronheim 2000)	Serious[1]	N/A	Not serious	Serious[2]	99	RR 9.55 (0.53, 172.81)	Low
Decision to forgo antibiotics							
1 (Ahronheim 2000)	Serious[1]	N/A	Not serious	Serious[2]	99	RR 7.43 (0.39, 140.15)	Low
Death in hospital							
1 (Ahronheim 2000)	Serious[1]	N/A	Not serious	Serious[2]	99	RR 1.06 (0.53, 2.13)	Low
Hospital admissions							
1 (Ahronheim 2000)	Serious[1]	N/A	Not serious	Serious[2]	99	MD 0.04 (-0.74, 0.82)	Low
New feeding tube							
1 (Ahronheim 2000)	Serious[1]	N/A	Not serious	Serious[2]	99	RR 1.06 (0.68, 1.65)	Low
Total feeding tube use							
1 (Ahronheim 2000)	Serious[1]	N/A	Not serious	Serious[2]	99	RR 1.06 (0.81, 1.39)	Low
Mechanical ventilation							
1 (Ahronheim 2000)	Serious[1]	N/A	Not serious	Serious[2]	99	RR 0.53 (0.10, 2.77)	Low
Tracheostomy							
1 (Ahronheim 2000)	Serious[1]	N/A	Not serious	Serious[2]	99	RR 0.35 (0.01, 8.84)	Low
Cardiopulmonary resuscitation							
1 (Ahronheim 2000)	Serious[1]	N/A	Not serious	Serious[2]	99	RR 0.15 (0.01, 2.86)	Low

1. Allocation assignment unclear and participants not blinded.
2. Non-significant result.

Use of decision aid on feeding options

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
Decisional conflict in surrogate decision-makers							
1 (Hanson 2011)	Serious[1]	N/A	Not serious	Serious[2]	90	MD -0.30 (-0.61, 0.01)	Low
Feeding discussion with physician, nurse practitioners or physician assistants							
1 (Hanson 2011)	Serious[1]	N/A	Not serious	Serious[2]	90	RR 1.57 (0.93, 2.64)	Low
Feeding discussion with other nursing home staff							

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
1 (Hanson 2011)	Serious[1]	N/A	Not serious	Serious[2]	90	RR 1.12 (0.86, 1.45)	Low
Any modified diet							
1 (Hanson 2011)	Serious[1]	N/A	Not serious	Serious[2]	90	RR 1.19 (0.31, 4.54)	Low
Specialised dysphagia diet							
1 (Hanson 2011)	Serious[1]	N/A	Not serious	Not serious	90	RR 1.30 (1.09, 1.56)	Moderate
Specialised staff assistance							
1 (Hanson 2011)	Serious[1]	N/A	Not serious	Serious[2]	90	RR 2.39 (0.81, 7.07)	Low
Specialised utensils							
1 (Hanson 2011)	Serious[1]	N/A	Not serious	Serious[2]	90	RR 0.24 (0.03, 2.06)	Low
Head/body positioning							
1 (Hanson 2011)	Serious[1]	N/A	Not serious	Serious[2]	90	RR 2.87 (0.12, 68.60)	Low

1. Participants and assessors not blinded.
2. Non-significant result.

Goals of Care intervention versus usual care

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
Quality of communication (overall) – higher numbers favour intervention							
1 (Hanson 2017)	Serious[1]	N/A	Not serious	Serious[2]	299	MD 0.20 (-0.29, 0.69)	Low
Quality of communication (general) – higher numbers favour intervention							
1 (Hanson 2017)	Serious[1]	N/A	Not serious	Serious[2]	299	MD 0.40 (-0.08, 0.88)	Low
Quality of communication (end of life) – higher numbers favour intervention							
1 (Hanson 2017)	Serious[1]	N/A	Not serious	Not serious	299	MD 0.80 (0.15, 1.45)	Moderate
Family-care provider concordance on primary care goal – higher numbers favour intervention							
1 (Hanson 2017)	Serious[1]	N/A	Not serious	Not serious	299	RR 1.24 (1.11, 1.40)	Moderate
Advanced care planning problem score >1 – lower numbers favour intervention							
1 (Hanson 2017)	Serious[1]	N/A	Not serious	Serious[2]	299	RR 1.03 (0.88, 1.20)	Low
Symptom management – higher numbers favour intervention							

Number of RCTs	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
1 (Hanson 2017)	Serious[1]	N/A	Not serious	Serious[2]	299	MD -1.10 (-3.18, 0.98)	Low
Satisfaction with care – higher numbers favour intervention							
1 (Hanson 2017)	Serious[1]	N/A	Not serious	Serious[2]	299	MD -0.60 (-1.87, 0.67)	Low
Palliative care treatment plan domain score – higher numbers favour intervention							
1 (Hanson 2017)	Serious[1]	N/A	Not serious	Not serious	299	MD 0.60 (0.13, 1.07)	Moderate

1. Participants not blinded.
2. Non-significant result.

Enteral tube feeding

Number of studies	Risk of bias	Inconsistency	Indirectness	Imprecision	Sample size	Effect size (95% CI)	Quality
Systematic review of enteral tube feeding studies							
Sampson (2009)	Serious[1]	N/A	Not serious	Serious[2]	1,813	No meaningful effects identified	Low

1. All included studies were observational studies at high risk of bias, but risk of bias upgraded from very serious to serious due to large sample size and consistent results
2. No meaningful differences identified between groups.