

**The Journeying through Dementia psychosocial intervention versus usual care study: a single-blind, parallel group, phase 3 trial**

MOUNTAIN, GA, COOPER, CL, WRIGHT, J, WALTERS, SJ, LEE, E, CRAIG, Claire <<http://orcid.org/0000-0002-3475-3292>>, BERRY, K, SPRANGE, K, YOUNG, T, MONIZ-COOK, E, DENING, T, LOBAN, A, TURTON, E, BERESFORD-DENT, J, THOMAS, BD, THOMPSON, BJ and YOUNG, EL

Available from Sheffield Hallam University Research Archive (SHURA) at:  
<https://shura.shu.ac.uk/30091/>

---

This document is the Supplemental Material

**Citation:**

MOUNTAIN, GA, COOPER, CL, WRIGHT, J, WALTERS, SJ, LEE, E, CRAIG, Claire, BERRY, K, SPRANGE, K, YOUNG, T, MONIZ-COOK, E, DENING, T, LOBAN, A, TURTON, E, BERESFORD-DENT, J, THOMAS, BD, THOMPSON, BJ and YOUNG, EL (2022). The Journeying through Dementia psychosocial intervention versus usual care study: a single-blind, parallel group, phase 3 trial. *The Lancet Healthy Longevity*, 3 (4), e276-e285. [Article]

---

**Copyright and re-use policy**

See <http://shura.shu.ac.uk/information.html>

# THE LANCET

## Healthy Longevity

### **Supplementary appendix**

This appendix formed part of the original submission and has been peer reviewed.  
We post it as supplied by the authors.

Supplement to: Mountain GA, Cooper CL, Wright J, et al. The Journeying through Dementia psychosocial intervention versus usual care study: a single-blind, parallel group, phase 3 trial. *Lancet Healthy Longev* 2022; **3**: e276–85.

## Appendix.

### 1. Additional statistical analysis for Journeying through Dementia Randomised Controlled Trial, Resource use costs and sources and outcome measures with additional references for measures not provided in the paper.

#### Intervention details and attendance

Table 1 below shows a summary of the JtD intervention and participant attendance. There were 28 intervention groups across the 13 sites. The median intervention size was five participants (IQR 4 to 6). Of the 241 participants randomised to receive JtD intervention, 165 attended at least 10 of the 16 available sessions (defined in the protocol as the per-protocol treatment threshold). The median number of group sessions attended (out of a possible 12) was 10 (IQR 1 to 11) and 153 out of a possible 244 participants attended all four individual meetings. Of those that attended at least one group session but less than 12, the majority (n=130) had intermittent session non-attendance, and a minority (n=23). having missed a session, did not attend future sessions.

**Table 1: Summary of attendance at JtD intervention**

Intervention	Summary
Number of JtD intervention groups	28
Session size	
N (total number of possible sessions)	336
Mean (SD)	5.3 (1.8)
Median (IQR)	5 (4, 6)
Min., Max.	1, 9
Meetings attended (out of a possible 16)	
N	244
Mean (SD)	10 (6)
Median (IQR)	13 (3, 15)
Min., Max.	0, 16
Did not attend intervention, n (%)	28 (11%)
Attended between 1 and 9 sessions, n (%)	48 (20%)
Attended between 10 and 16 sessions (per-protocol), n (%)	168 (69%)*
Group meetings attended (out of a possible 12)	
N	244
Mean (SD)	7.3 (4.6)
Median (IQR)	9 (1, 11)
Min., Max.	0, 12
Individual meetings attended, n(%)	
0	28 (11%)
1	33 (14%)
2	12 (5%)
3	18 (7%)
4	153 (63%)

Meetings accompanied, n(%)	Group	Individual
0	102 (42%)	75 (31%)
1	47 (19%)	42 (17%)
2	45 (18%)	38 (16%)
3	50 (20%)	30 (12%)
4		59 (24%)

Accompanied by, n(%)	
Participating supporter	523 (78%)
Non-participating supporter	32 (5%)
Consultee	9 (1%)
Other	108 (16%)

The number that received at least 10 sessions is higher than that shown in the CONSORT flow diagram as it includes three control group participants that received and complied with the intervention in error. Accompanied by data was not collected for some participants within the group sessions. Efforts were made to retrospectively verify who accompanied participants, but facilitators did not always hold this information.

### Demographics of intervention attendees

Table 2 shows intervention attendance by demographic and socioeconomic variables of interest. Intervention ‘attenders’ (those attending at least 10 of a possible 16 sessions) were more likely to be male (61% versus 48% in the ‘non-attenders’) and have a participating supporter (80% versus 58%). Participants’ age, type and length of dementia diagnosis, and socio-economic status (based on Index of Multiple Deprivation decile) were similar for those that attended at least 10 sessions and those that attended less than 10 sessions in the intervention arm.

**Table 2: Participant characteristics by intervention attendance status for participants randomised to the intervention arm (n=241) (table 13 report)**

Characteristic	Intervention Attendance		
	At least 10 sessions (n=165)	less than 10 session (n=76)	All (n=241)
Sex			
Male	100 (61%)	36 (47%)	136 (56%)
Female	65 (39%)	40 (53%)	105 (44%)
Participating supporter			
No	32 (19%)	32 (42%)	64 (27%)
Yes	133 (81%)	44 (58%)	177 (73%)
Type of dementia diagnosed			
Alzheimer’s	101 (61%)	41 (54%)	142 (59%)
Vascular dementia	19 (12%)	12 (16%)	31 (13%)
Mixed Alzheimer’s / vascular dementia	32 (19%)	19 (25%)	51 (21%)
Dementia in Parkinson disease	3 (2%)	0 (0%)	3 (1%)
Frontotemporal dementia (FTD)	4 (2%)	1 (1%)	5 (2%)
Lewy body dementia	1 (1%)	0 (0%)	1 (0%)
Unspecified dementia	5 (3%)	2 (3%)	7 (3%)
Other	0 (0%)	1 (1%)	1 (0%)

Age (years)			
N (%)	165 (100%)	76 (100%)	241 (100%)
Mean (SD)	77.2 (6.9)	78.2 (7.1)	77.5 (7.0)
Median (IQR)	78.0 (73.0, 81.0)	78.5 (73.5, 83.0)	78.0 (73.0, 82.0)
	56.0, 93.0	58.0, 91.0	56.0, 93.0
Index of Multiple Deprivation Decile			
N (%)	164 (99%)	76 (100%)	240 (100%)
Mean (SD)	6.3 (2.9)	5.9 (3.1)	6.2 (3.0)
Median (IQR)	7.0 (4.0, 9.0)	6.0 (3.0, 9.0)	7.0 (4.0, 9.0)
Min., Max.	1.0, 10.0	1.0, 10.0	1.0, 10.0
Length of time since dementia diagnosis (yrs)	165 (100%)	76 (100%)	241 (100%)
N (%)	1.3 (1.4)	1.3 (1.5)	1.3 (1.5)
Mean (SD)	0.8 (0.3, 1.8)	0.6 (0.2, 1.9)	0.7 (0.3, 1.8)
Median (IQR)	0.0, 7.7	0.0, 7.9	0.0, 7.9

## Measures taken

The measures taken with both the participant and participating carer at each of the visits are detailed in Tables 3 and 4.

**Table 3: Assessments completed with the participant**

Measure	Baseline Visit	8-Month Follow-Up Visit	12-Month Follow-Up Visit
DEMQL	X	X	X
EQ-5D-5L	X	X	X
PHQ-9	X	X	
GAD-7	X	X	
GSE	X	X	
Diener's Flourishing Scale	X	X	
SMAS	X	X	
IADL	X	X	
HSCRU		X	X

Mini Mental State Examination (MMSE) is measured on a scale from 0-30, higher scores indicate better cognitive function; we used the cut-off score of 18 or over for mild dementia to identify the trial population.

Dementia Related Quality of Life (DEMQL) is measured on a scale from 28 to 112, higher scores represent higher health related QoL.

Patient Health Questionnaire (PHQ-9)

is measured on a scale from 0 to 27, with higher scores indicating more severe depressive symptoms.

Generalised Anxiety Disorder-7 (GAD-7) is measured on a scale from 0 to 21, higher scores represent increasing severity of anxiety.

European Quality of Life – 5 Dimensions version (EQ-5D-5L) score is measured on a scale from –0.224 to 1.00 (full health).

EQ-5D VAS is measured on a scale from 0 (worst imaginable health state) to 100 (best imaginable health state).

General Self-Efficacy scale (GSE) is measured on a scale from 10 to 40, higher scores indicate more self-efficacy.

Diener's Flourishing scale is measured from 0 to 56, higher scores represent more psychological resources and strengths.

Self-Management Assessment Scale (SMAS) is measured on a scale from 30 to 175, higher score indicates greater self management ability. (1)

Instrumental Activities of Daily Living (IADL) is measured on a scale from 0 to 8, higher scores represent lower level of dependence.

Lawton MP, Brody EM. Assessment of Older People: Self-Maintaining and Instrumental Activities of Daily Living 1. (2)

Health and Social Care Resource Use (HSCRU). See Tables 10- 15 below.

**Table 4: Assessments completed with the participating supporter**

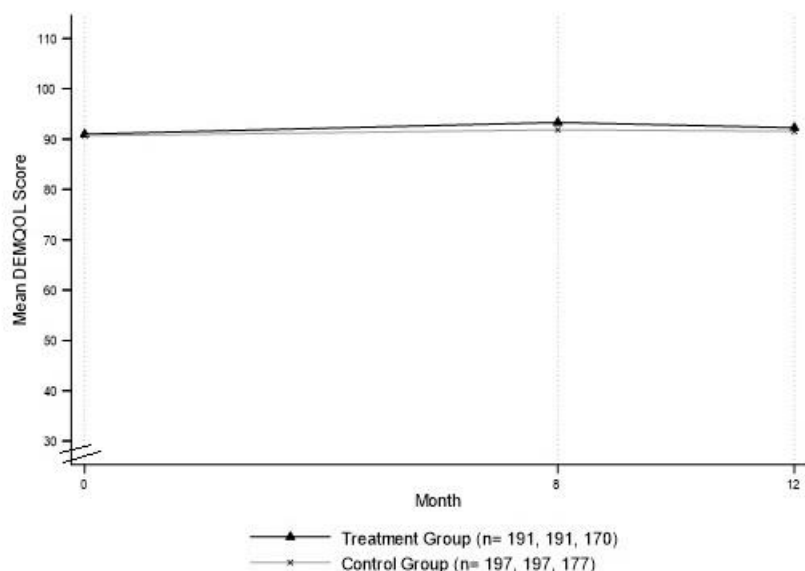
Measure	Baseline Visit	8-Month Follow-Up Visit	12-Month Follow-Up Visit
EQ-5D-5L	X	X	
PHQ-9	X	X	
SCQ	X	X	

Sense of Competence questionnaire (SCQ) is measured on a scale of 27 to 135, a higher score represents a person with a better sense of competence.

### Statistical analysis

Figure 1 shows the unadjusted mean DEMQOL score (primary outcome) by timepoint and allocated group, the difference between the treatments at 12 months was smaller still than at 8 months.

**Figure 1: Unadjusted mean DEMQOL score (primary outcome) by timepoint and allocated group**



A number of sensitivity analyses were undertaken on the 8 month DEMQOL score with regard to alternative missing data assumptions, removing mis-timed measurements and treatment adherence. In all cases the treatment difference remained similar, partly due to drop out being low and adherence being high and so the populations included in the sensitivity analysis only differ slightly to those included in the primary analysis. Complier Average Causal Effects (CACE) analysis estimated a treatment difference as slightly larger than the ITT analysis (1.5 points) but the 95% CI for the treatment difference was still below 4 points (-0.6 to 3.7; p=0.158)

**Table 5: Sensitivity analysis results for the comparison of mean 8-month DEMQOL score between randomised treatment groups**

Outcome - DEMQOL 8months	Treatment group				Adjusted <sup>a</sup>		
	Intervention	Control	Intervention	Control	mean difference	95% CI	p-value
	N	Mean (SD)	n	Mean (SD)			
Observed Data (primary analysis)	191	93.3 (13.0)	197	91.9 (14.6)	0.9	-1.2 to 3.0	0.38
Removing mis-timed measurements <sup>b</sup>	182	93.3 (13.1)	188	92.0 (14.5)	1	-1.1 to 3.1	0.37
Simple Regression imputation <sup>c</sup>	235	93.3 (12.3)	236	91.8 (13.9)	0.8	-0.9 to 2.6	0.33
Multiple imputation <sup>d</sup>	235	93.3 (13.0)	236	91.8 (14.6)	0.8	-1.3 to 2.9	0.47
Per-protocol <sup>e</sup>	155	93.7 (13.2)	197	91.9 (14.6)	1.2	-0.9 to 3.3	0.26
CACE <sup>f</sup>					1.5	-0.6 to 3.7	0.16

DEMQOL is measured on a scale from 28 to 112, higher scores represent higher health related QoL. A positive mean difference implies the Intervention group had the better health related quality of life.

- a) Adjusted for baseline DEMQOL score, stratification site and JtD intervention group
- b) removed outcome measures taken outside the window of 2 weeks before to 8 weeks after 8-month follow up was due.
- c) regression imputation used sex, age, presence of supporter, type of dementia and baseline DEMQOL, PHQ-9 and GAD-7 scores as covariates.
- d) Multiple imputation using chained equations (regression) based on 100 imputed data sets, with age, sex, stratification site, presence of supporter, type of dementia, duration of dementia and baseline DEMQOL, GAD-7 and PHQ-9 score as covariates.
- e) defined as attending at least 10 of the 16 possible sessions.
- f) Complier Average Causal Effect (CACE) using two stage least squares regression with age, sex, presence of supporter, stratification site and baseline DEMQOL as covariates and standard errors that allow for intragroup correlation by JtD intervention groups. All other analyses use a mixed effects regression model, with clustering in the intervention arm only.

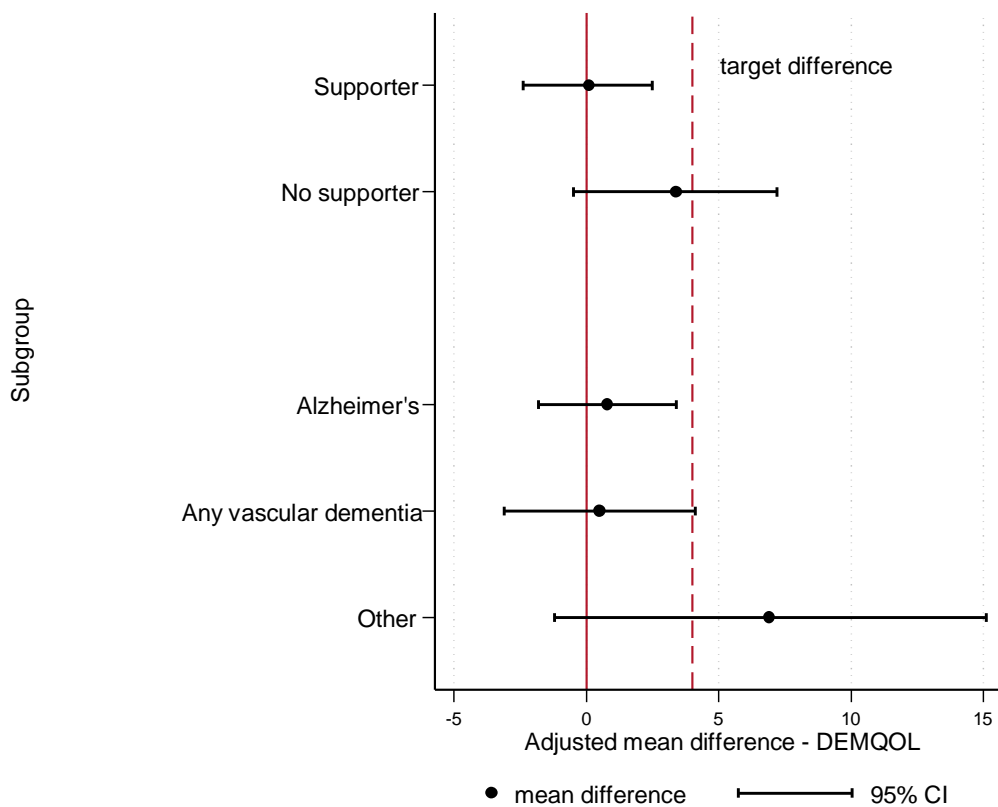
The potential moderating effect of two predefined subgroups were explored by including an interaction between treatment and subgroup in the mixed effects primary outcome regression model as shown in table 6 and summarised in figure 2. No reliable statistical evidence of subgroup effects or interactions were found between the treatment groups.

**Table 6: Comparison of mean 8-month DEMQOL score between treatment group by subgroup**

Outcome	Treatment group		Control		mean difference	95% CI	interaction coefficient <sup>a</sup>	95% CI	p-value
	Intervention								
	n	Mean (SD)	n	Mean (SD)					
Is there a participating supporter?									
Yes	145	93.6 (13.2)	139	92.8 (14.3)	0.1	-2.4 to 2.5			
No	46	92.5 (12.4)	58	89.7 (15.2)	3.4	-0.5 to 7.2	3.3 <sup>b</sup>	-1.2 to 7.7	0.15
Type of dementia									
Alzheimer's	117	94.1 (14.2)	126	93.4 (14.0)	0.8	-1.8 to 3.4			
Any vascular dementia	58	92.7 (9.9)	62	91.1 (14.8)	0.5	-3.1 to 4.1			
Other	16	90.1 (14.5)	9	76.5 (13.8)	6.9	-1.2 to 15.1			0.34
Type of dementia - sensitivity 1									
Alzheimer's	117	94.1 (14.2)	126	93.4 (14.0)	0.8	-1.8 to 3.4			
Vascular dementia	20	91.3 (10.5)	14	89.5 (11.4)	2.8	-3.9 to 9.5			
Mixed and other dementia	54	92.5 (11.3)	57	89.2 (16.3)	1	-2.8 to 4.8			0.86
Type of dementia - sensitivity 2									
Alzheimer's and mixed dementia	155	93.9 (13.2)	174	92.9 (14.4)	0.6	-1.7 to 2.8			
Vascular dementia	20	91.3 (10.5)	14	89.5 (11.4)	2.8	-3.9 to 9.5			
Other dementia	16	90.1 (14.5)	9	76.5 (13.8)	7	-1.1 to 15.2			0.28

a) Adjusted for baseline score, stratification site and JtD intervention group. b. interaction coefficient for intervention x no participating supporter (additional effect of having intervention if no supporter compared to supporter).

**Figure 2: Mean 8-month DEMQOL score by subgroup and treatment group (n=388) (figure 11 report)**



## Serious adverse events

**Table 7: Serious adverse events by treatment group (n=480) (table 21 report)**

Serious Adverse Events	Intervention (n=241)	Control (n=239)	All (n=480)
Number of participants who experienced $\geq 1$ SAE	40	35	75
Number of all SAEs (including repeated events)	61	39	100
Occurred			
Before randomisation	3 (5% <sup>a</sup> )	1 (3%)	4 (4%)
After randomisation	58 (95%)	38 (97%)	96 (96%)
Seriousness			
Death	10 (16%)	5 (13%)	15 (15%)
Life threatening	3 (5%)	2 (5%)	5 (5%)
Inpatient hospitalisation	45 (74%)	31 (79%)	76 (76%)
Prolongs hospitalisation	1 (2%)	0 (0%)	1 (1%)
Persistent or significant disability/incapacity	2 (3%)	1 (3%)	3 (3%)
Intensity <sup>b</sup>			
Mild	6 (10%)	5 (13%)	11 (11%)

Moderate	35 (57%)	25 (64%)	60 (60%)
Severe	17 (28%)	9 (23%)	26 (26%)
Relationship to intervention			
Unlikely	13 (21%)	5 (13%)	18 (18%)
Unrelated	48 (79%)	34 (87%)	82 (82%)

Two SAEs were recorded on participants who withdrew before randomisation and are not included in the table. a) The denominator for all percentages is the number of SAEs

b) Information on the intensity of three SAEs was not available due to: 1) inability to access patient medical records after withdrawal, 2) PI not being able to assess intensity based on available data, 3) data being missing.

### Participating carers

Table 8 gives the baseline characteristics of participating carers; the majority were female (74%) and were the spouse/ partner of the participant (79%)

**Table 8: Baseline characteristics and participant reported outcome measures by randomised treatment group for all participating carers of randomised participants (n=350)**

Characteristic		Intervention (n=177)	Control (n=173)	All (n=350)
Age	N (%)	171 (97%)	166 (96%)	337 (96%)
	Mean (SD)	70 (11)	68 (12)	69 (11)
	Median (IQR)	73 (65, 77)	70 (60, 76)	71 (62, 76)
	Min., Max.	37, 89	26, 88	26, 89
Sex		52 (29%)	39 (23%)	91 (26%)
	Male	125 (71%)	134 (77%)	259 (74%)
	Female	52 (29%)	39 (23%)	91 (26%)
Relationship to person with dementia	Spouse/Partner	140 (79%)	136 (79%)	276 (79%)
	Child	31 (18%)	32 (18%)	63 (18%)
	Sibling	1 (1%)	1 (1%)	2 (1%)
	Other family member	3 (2%)	2 (1%)	5 (1%)
	Friend	0 (0%)	2 (1%)	2 (1%)
	Neighbour	1 (1%)	0 (0%)	1 (0%)
	Other	1 (1%)	0 (0%)	1 (0%)
Length of time caring for person with dementia*	0-6 months	35 (20%)	24 (14%)	59 (17%)
	7-12 months	19 (11%)	12 (7%)	31 (9%)
	13-18 months	18 (10%)	20 (12%)	38 (11%)
	19-24 months	10 (6%)	19 (11%)	29 (8%)
	> 2 years	92 (52%)	97 (56%)	189 (54%)
Lives with person with dementia	No	26 (15%)	31 (18%)	57 (16%)
	Yes	151 (85%)	142 (82%)	293 (84%)

EQ-5D-5L (crosswalk value index)	Mean (SD)	0.81 (0.18)	0.82 (0.18)	0.81 (0.18)
	N (%)	168 (95%)	168 (97%)	336 (96%)
EQ-5D VAS	Mean (SD)	77.8 (18.9)	79.7 (16.2)	78.8 (17.6)
	N (%)	169 (95%)	168 (97%)	337 (96%)
SCQ (total score)	Mean (SD)	100.3 (16.7)	102.5 (16.3)	101.4 (16.6)
	N (%)	159 (90%)	160 (92%)	319 (91%)
PHQ-9 (total score)	Mean (SD)	4.1 (4.2)	3.9 (4.4)	4.0 (4.3)
	N (%)	167 (94%)	168 (97%)	335 (96%)

EQ-5D-5L score is measured on a scale from -0.224 to 1.00 (full health).

EQ-5D VAS is measured on a scale from 0 (worst imaginable health state) to 100 (best imaginable health state).

SCQ is measured on a scale from 27 to 135, higher score represents a person with a better sense of competence.

PHQ-9 is measured on a scale from 0 to 27, with higher scores indicating more severe depressive symptoms.

\* Responses were missing from participating carers in both the intervention and control arms.

As shown in table 9, The mean differences in 8-month outcome measures for the participating carers are small and non-significant with the exception of the EQ-5D where the carers in the control group reported better health (adjusted mean difference -0.06, 95% CI to -0.02; p=0.002). Such a finding has been reported in other studies where family carers are asked to undertake additional activity. (3) The difference was not borne out in the EQ-5D VAS which showed no difference between the intervention and control groups (mean difference 0.1; 95% CI -2.8 to 3.1).

**Table 9: comparison of mean 8-month secondary quality of life outcomes for the participating carers (n=274)**

Outcome	Treatment group				Adjusted*		
	Intervention		Control				
	N	Mean (SD)	N	Mean (SD)	mean difference	95% CI	p-value
EQ-5D-5L (crosswalk value index)	138	0.76 (0.22)	135	0.80 (0.17)	-0.06	-0.09 to -0.02	0.002
EQ-5D VAS	138	77.9 (15.5)	136	77.7 (16.8)	0.1	-2.8 to 3.1	0.94
SCQ (total score)	128	97.8 (17.3)	125	101.3 (16.6)	-1.4	-4.3 to 1.5	0.34
PHQ-9 (total score)	139	4.7 (4.6)	135	4.1 (4.6)	0.4	-0.4 to 1.2	0.35

\*Adjusted for baseline score and stratification site. EQ-5D-5L score is measured on a scale from -0.224 to 1.00 (full health). EQ-5D VAS is measured on a scale from 0 (worst imaginable health state) to 100 (best imaginable health state). SCQ is measured on a scale from 27 to 135, higher score represents a person with a better sense of competence. PHQ-9 is measured on a scale from 0 to 27, with higher scores indicating more severe depressive symptoms. For the EQ-5D and SCQ outcomes a positive mean difference implies the Intervention group had the better health-related quality of life. For the PHQ-9 outcome a positive mean difference implies the Intervention group had worse health-related quality of life.

## Health economics analysis

### Resource use cost sources

1. NHS Improvement Reference Costs 2017/18 [Internet] Available from: <https://improvement.nhs.uk/resources/reference-costs/> (accessed 15 Jan 2019)
2. Curtis L, Burns A. Unit Costs of Health & Social Care 2018. Canterbury: University of Kent; 2018
3. Royal College of Nursing. NHS pay scales 2017-18 [Internet]. RCN. 2019. Available from: <https://www.rcn.org.uk/employment-and-pay/nhs-pay-scales-2017-18> (accessed 16 Dec 2019).
4. The National Institute for Health and Care Excellence (NICE). British National Formulary [Internet]. NICE. 2019. Available from: <https://bnf.nice.org.uk/> (accessed 16 Dec 2019).

**Table 10: Number of and reasons for out-patient visits and associated by randomised group**

Speciality	Source	Unit Cost	JtD	Usual Care	Total
Anticoagulation	DoH OP 324	£34	1	1	2
Audiology	DoH OP 310	£108	9	6	15
Brain scan	DoH IMAGOP	£141	1	0	1
Breast Screening		£12	0	1	1
Plastic surgery	DOH OP 104	£150	1	0	1
Burns unit	DoH OP	£156	1	0	1
Cardiology	DoH OP 320	£134	13	10	23
Chiropodist	PSSRU	£44	1	0	1
General out-patient	DoH OP	£125	1	7	8
CT	DoH IMAGOP	£228	1	0	1
Dental	DoH OP 141	£152	1	0	1
Dermatology	DoH OP 330	£111	9	8	17
Diabetes	DoH OP	£147	0	4	4
Dialysis	DoH RENAL CKD	£151	0	3	3
Dietician	PSSRU T7.1	£86	0	1	1
Endocrinology	DoH OP 302	£164	0	2	2
ENT	DoH OP 120	£104	4	3	7
Epidural	DoH	£412	1	0	1
Gastroenterology	DoH OP 301	£149	8	6	14
Geriatric Outpatients	DoH OP 430	£257	1	0	1
GI Physiology	DoH OP 106	£147	1	0	1
Gynaecology	DoH OP 502	£142	0	2	2
Haematology	DoH OP 303	£160	4	6	10
Limb fitting	DoH OP	£142	0	1	1
Maxillofacial Surgery	DoH OP 144	£126	1	0	1
Memory clinic	DoH OP 727	£136	2	5	7
Neurology	DoH OP 400	£167	8	9	17
Occupational Health	DoH OP 651	£73	2	0	2
Oncology	DoH OP 370	£162	7	3	6
Ophthalmology	DoH OP 130	£98	15	23	38
Orthopaedics	DoH OP 111	£124	4	7	11
Orthotics	DoH OP 658	£126	3	0	3
Outpatient Nurse	PSSRU	£24	1	0	1
Pain Clinic	DoH OP	£145	1	2	3
Phlebotomy	DoH OP	£34	2	1	3
Physiotherapy	DoH OP 304	£74	1	3	4
Plastics	DoH OP	£107	0	1	1
Podiatry	DoH OP 653	£51	3	2	5
Pulmonology	DoH OP	£150	0	1	1
Radiology	DoH OP	£145	6	8	14
Respiratory	DoH OP	£24	2	2	4
Rheumatology	DoH OP 410	£146	2	6	8
Scan	DoH OP	£141	1	0	1
Speech and Language Therapy	PSSRU	£55	1	0	1
Stroke/seizures	DoH OP 328	£216	1	0	1
Surgical	DoH OP	£140	1	0	1
Ultrasound	DoHIMAGOP	£54	0	1	1
Urgent care	DoH UC	£160	1	0	1
Urology	DoH OP 107	£110	9	8	17
Vascular	DoH OP	£147	1	1	2
Walk-in Clinic	PSSRU GP	£31	0	1	1
X ray	DoH IMAGOP	£71	3	3	5
Total			136	153	289

**Table 11: Reason for A&E visits and number of visits by randomised group**

	JtD	Usual Care	Total
Increased confusion	1	0	1
Ankle	1	0	1
Breathing problems	2	2	4

Broken Hip	0	1	1
COPD	1	0	1
Cancer - acute pain	1	0	1
Chest Pain	1	1	1
Collapsed	1	0	1
Fall	9	2	11
Grazed leg	1	0	1
Haemorrhage	1	0	1
Heart attack	0	1	1
INR excessively high	1	0	1
Inflamed tendon and displaced Achilles	1	0	1
Pain compacted bowel	0	1	1
Pain in back	0	1	1
Pain in chest/abdominal	0	1	1
Pain in leg	1	0	1
Pneumonia	0	1	1
Problems with Catheter	0	2	2
Severe constipation	0	1	1
Suspected atypical TIA	0	1	1
Swollen leg	1	0	1
Dog bite	0	1	1
Urinary issues	0	2	2
Total	23	18	41

**Table 12: Resource use details, unit costs and sources of unit costs for A&E visits, Emergency & Non-Emergency transport, lunch clubs and other social clubs**

	Source	Unit Cost	JtD	Usual Care
A&E visits:	DoH reference costs AE	£160		
0			144	181
1			19	14
2			2	2
Emergency Transport:	DoH reference cost	£222		
0	(mean of see treat or		157	187
1	refer and see treat and		1	9
2	convey)		7	1
Non-emergency transport	Dialysis Transport:	£22		
0	Finding a way together.			
1	Kidney Care UK 2019		158	194
2	Page 20		4	2
26*			3	0
Day care	PSSRU Table 1.4 (page 30)	£58	2	2
Lunch clubs, social clubs and other clubs	PSSRU Table 1.4 (page 30)	£13 per hour	17	12

\* Person on dialysis

**Table 13: Number of Community Services visits/contacts and unit cost of contact**

	Source	Unit Cost	JtD	Control	Overall
GP	PSSRU	£31	157	161	318
Practice Nurse	PSSRU Table 10.2	£36	76	81	112
Memory Clinic	Reference cost Out Patient	£136	26	21	47
Chiropodist	PSSRU Band 5	£34	21	22	43
District Nurse	PSSRU	£53	13	17	30
Social Worker	PSSRU	£44	8	14	22
Physiotherapist	PSSRU Table 17.1	£46	6	15	21
Home Care Worker	PSSRU	£22	9	8	17
Community MH nurse	PSSRU per contact	£9	4	12	16
Occupational therapy	PSSRU Band 5	£34	10	3	13
Dentist	PSSRU Table 10.6	£104	3	7	10
Nurse	PSSRU Nurse Band 5	£37	4	4	8
Diabetes Nurse	PSSRU Hospital Nurse	£54	1	4	5
Pharmacist	PSSRU Band 6	£45	3	2	5
Speech and language	PSSRU Band 5	£34	3	2	5
Dietician	PSSRU Table 7.1	£86	2	2	4
Optician	PSSRU	£31	3	1	4
Podiatrist	PSSRU Band 5	£34	4	0	4
Psychologist	PSSRU Band 7	£54	2	2	4
Phlebotomist	PSSRU Mean cost per contact	£0.48	1	2	3
Audiology	Reference cost Out Patient	£108	1	1	2
Community Psychiatrist	PSSRU per contact	£9.43	6	1	2
First responder	Assume same as paramedic PSSRU	£52	1	1	2
Heart Nurse	PSSRU Band 7	£53	0	2	2
Balancing Group	Reference cost Out Patient	£104	0	1	1
Cardiac Nurse	PSSRU Band 7	£53	0	1	1

Care Allowance Assessment	Assume social worker per hour	£44	0	1	1
	PSSRU				
Carers	PSSRU per hour	£22	1	0	1
Cognitive Stimulation	PSSRU Table 6.7	£52	1	0	1
Counsellor	PSSRU Band 5	£36	2	0	1
CST	PSSRU Table 6.7	£52	0	1	1
Health Visitor	Assume the same as home care worker PSSRU	£22	0	1	1
Home assessment	Assume the same as home care worker PSSRU	£22	0	1	1
Hospice	PSSRU Table 7.6 per hour of service	£76	0	1	1
Integrated care centre	Costed as social worker PSSRU	£44	0	1	1
Paramedic	Assume part of ambulance as also ticked ambulance required		0	1	1
Psychiatrist	PSSRU	£341	0	1	1
Sensor team	Costed as social worker PSSRU	£44	0	1	1
Thyroid Clinic	PSSRU Nurse Band 5	£37	1	0	1
Walk In Centre	<b>PSSRU (Assume same as GP)</b>	<b>£31</b>	0	1	1
Warfarin clinic	PSSRU Nurse band 5	£37	0	1	1

**Table 14: Number of Hospitalisations by reason for admission, average length of stay (LOS) of hospitalisation and unit cost of hospitalisation**

Reason for admission	Source <sub>1</sub>	Unit Cost	Average LOS	JtD	Usual Care
Abdominal pain	Elective Inpatient FD05A	£3075	4	1	0
Acute neuropathic pain	Elective Inpatient WH08A	£2,063	5	0	1
Breathing problems	Elective Inpatient DZ18D	£1351	2	0	2
Broken Hip	Elective Inpatient HE11C	£9733	14	0	1
COPD	Elective Inpatient DZ65E	£3737	5	1	0
Chest pains	Elective Inpatient EB12B	£728	1	1	0
Back issues	Elective Inpatient HC32G	£2526	2	1	0
Detox (alcohol)	Elective Inpatient WH21B	£495	1	0	1
Diarrhoea/dehydration	Elective Inpatient FD01D	£3747	6	1	0
Endoscopy/Investigation	Elective Inpatient FE12A	£1015	1	1	0
Haemorrhage	Elective Inpatient AA23F	£2367	3	1	0
Heart attack	Elective Inpatient EB10A	£3940	5	0	1
Increased confusion ?UTI	Elective Inpatient LA04K	£4487	11	1	0
Pneumonia	Elective Inpatient DZ11P	£4993	10	1	1
Stroke	Elective Inpatient AA35C	£4511	11	0	1
Suspected atypical TIA	Elective Inpatient AA29D	£1109	2	0	1
Urinary issues	Elective Inpatient LA04K	£4487	11	0	2
Compacted bowel	Elective Inpatient FD02C	£5447	10	0	1
Prolapse operation	Elective Inpatient MB09C	£3171	3	0	1
<b>Total</b>				<b>9</b>	<b>13</b>

Department of Health Reference Costs 2017-18

**Table 15: Other hospital visits**

Service used (e.g. day hospital, care home admission)	Source	Unit Cost	JtD	Usual Care
Community hospital	Reference cost Community Health Services	£296	1	0
Day Hospital	Reference costs	£327	1	5
Eye Clinic	Reference costs	£98	1	0
GP/Nurse	PSSRU	£31	1	1
Podiatry	Reference Cost Community Health Services	£53	0	1
Vascular surgery	Reference cost day case VQ50C	£481	1	0
Warfarin clinic	Reference cost consultant led pain management	£101	0	1
Phlebotomy	Reference cost	£34	1	0
Walk in centre	PSSRU (Assume same as GP)	£31	1	0
<b>Total</b>			<b>7</b>	<b>8</b>

**Table 16: Number and type of medications taken and unit cost/price of medication**

Medication	Route	Price	JtD	Usual Care
Allopurinol	Tablet	1.72	0	1
Alzain	Tablet	5.99	1	0

Alzest	Patch	19.97	3	4
Amitriptyline	Tablet	0.89	1	14
Amlodipine	Tablet	0.74	0	1
Ariprazole	Tablet	1.71	0	1
Aspirin	Tablet	0.83	2	4
Atenolol	Tablet	0.61	1	0
Atrovastatin	Tablet	1.20	1	0
Buprenorphine	Patch	12.62	2	3
	Tablet	6.34		
Buteac	Patch	22.12	0	3
Carbamazepine	Tablet	4.16	1	0
Certraline	Tablet	2.40	1	0
Circadin	Tablet	15.39	1	0
Citalopram	Tablet	0.83	3	28
Clomethiazole	Tablet	32.80	0	2
Clonazepam	Tablet	28.31	2	0
Co-codamol	Tablet	7.01	14	13
Co-dydromol	Tablet	7.25	2	0
Codeine	Tablet	0.98	11	3
Corticosteroid injection	Injection	1.16	0	1
Denepezil	Tablet	1.04	0	1
Diazepam	Tablet	0.59	3	5
Diconfenac	Gel	4.63	1	0
Dihydrocodeine				2
Donepezil	Tablet	1.56	155	173
Dosulepin	Tablet	2.97	0	2
Duloxetine	Tablet	2.95	2	5
Fenbid forte	Gel	2.45	4	1
Flucloxacillin	Tablet	2.48	0	1
Fluoxetine	Tablet	1.01	14	11
Gabapentin	Tablet	3.09	6	16
Galantamine	Tablet	55.16	14	8
Gatalin	Tablet	32.45	3	4
Glyceryl trinitrate	Spray	3.85	1	1
Ibuprofen	Tablet	0.91	9	9
Ibuprofen gel	Gel	5.79	1	1
Imipramine	Tablet	0.82	0	2
Lithodprom	Tablet	0.82	1	0
Lorazepam	Tablet	600.00	1	0
Luventa	Tablet	55.16	2	0
Madopar	Tablet	4.96	1	0
Melatonin	Tablet	65.00	1	2
Melaxican	Tablet	1.01	0	1
Memantine	Tablet	1.29	61	80
Methocarbamol	Tablet	12.66	0	1
Methotrexate	Injection	14.92	0	2
Mirtazapine	Tablet	1.97	8	24
Monomil xl	Tablet	3.49	0	1
Morphine	Patch/Liquid	19.45	2	3
Moveate cream	Cream	11.29	0	1
Mst continus	Tablet	3.29	1	2
Naproxen	Tablet	7.77	4	5
Nefopam	Tablet	7.77	0	2
Nortriptyline	Tablet	8.00	0	2
Nytol	Tablet	6.20	1	0
Omeprazole	Tablet	0.87	1	0
Oramorph	Oral	1.89	1	3
Oxycodone	Solution/Suppository	9.71	0	2
Paracetamol	Tablet	0.40	57	57
Paroxetine	Tablet	1.26	2	2
Pimozide	Tablet	40.31	0	1
Pregabalin	Tablet	4.99	1	5
Priadel	Tablet	2.76	1	0
Propranolol	Tablet	1.40	0	2
Quetiapine	Tablet	5.24	3	4
Quinine	Tablet	2.78	1	1
Rasgaline	Tablet	2.07	1	0
Resmed	Mask	101.00	0	2
Risperidone	Tablet	3.58	2	0
Rivastigmine	Tablet	5.95	12	11
	Patch	35.10	17	12
Rosuvastatin	Tablet	1.65	0	1
Sertraline	Tablet	1.27	35	22
Sinemet	Tablet	6.28	0	1
Sleepeze	Tablet	5.79	0	1
Steroid injection	Injection	1.16	1	0
Tegretol	Tablet	5.20	1	0
Temazepam	Tablet	1.56	0	2
Tetrabenazine	Tablet	100.00	0	1

Topiramate	Tablet	6.94	0	1
Tramadol	Tablet	4.60	13	10
Trazodone	Tablet	5.23	7	1
Voltarol	Gel	3.47	0	3
Zapain	Tablet	3.85	3	2
Zerobase	Cream	5.26	1	0
Zomorph	Tablet	8.30	1	5
Zopiclone	Tablet	1.19	5	7
<b>Total</b>			<b>548</b>	<b>605</b>

Unit cost source: All medication unit costs were obtained from the British National Formula (BNF) <https://bnf.nice.org.uk/>

### Sensitivity Analysis

Sensitivity analysis was applied to estimate utilities for the EQ-5D-5L values for the Devlin algorithm.(4) A preference-based version of the DEMQOL (DEMQOL-U) was also used in sensitivity analysis.(5) A further sensitivity analysis used multiple imputation to impute missing cost and QALY values; twenty imputations were run. (6)

Sensitivity analyses are presented in Table 17. Two further ways of estimating QALYs were explored but in both cases the utility loss was small and in favour of controls and the incremental cost-effectiveness of JtD remained dominated. Finally, missing QALYs and costs were imputed using multiple imputation, but JtD remained dominated by controls.

**Table 17. Mean costs and Quality adjusted life years (QALYs) for main analysis and sensitivity analysis with 95% Bias-Corrected Bootstrap Confidence Intervals**

	Intervention (N=166)	Control (N=173)	Incremental costs and QALYs	ICER*
<b>Main analysis</b>				
<b>QALYs</b>	<b>0.774</b> <b>(0.744, 0.802)</b>	<b>0.777</b> <b>(0.748, 0.803)</b>	<b>-0.003</b> <b>(-0.044, 0.038)</b>	
<b>Overall costs (Over 12 months)</b>	<b>£1,676</b> <b>(£1,367, £2,227)</b>	<b>£1,067</b> <b>(£792, £1,484)</b>	<b>£609</b> <b>(£105, £1,179)</b>	<b>-£202,857</b> <b>(-£534,733, £483,739)</b>
<b>Devlin algorithm(21)</b>				
QALYs	0.838 (0.813, 0.861)	0.841 (0.815, 0.864)	-0.003 (-0.038, 0.032)	
<b>DEMQOL-U(22)</b>				
QALYs	0.802 (0.789, 0.814)	0.818 (0.809, 0.828)	-0.016 (-0.032, -0.0004)	
Overall costs (Over 12 months)	£1,676 (£1,367, £2,227)	£1,067 (£792, £1,484)	£609 (£105, £1,179)	-£38,036 (-£224,443, £1,088)
<b>Imputing missing values</b>				
	N = 174	N = 176		
QALYs	0.768 (0.739, 0.796)	0.775 (0.745, 0.801)	-0.007 (-0.047, 0.033)	
Overall costs (Over 12 months)	£1,675 (£1,380, £2,256)	£1,058 (£790, £1,447)	£617 (£109, £1,183)	-£88,187 (-£475,761, £458,153)

\* Incremental cost-effectiveness ratio = incremental costs/incremental QALYs

A positive difference in incremental QALYs means the Intervention group had the better QALYs

A positive difference in incremental costs means the Intervention group had the higher cost.

**Table 18: Usual care treatments at JtD trial sites**

JtD Trial Site ID	Treatments provided as usual care						
	Conducting needs assessment	Drug therapy	Referral back to primary care	CST group	Educational group	Memory or 'Living well with dementia' group	Third sector */ Other **
S01	✓	✓	✓	✓	✓		
S02	✓	✓	✓				✓
S04	✓	✓			✓		✓
S05	✓	✓	✓	✓		✓***	✓
S06	✓	✓	✓	✓			
S07	✓	✓	✓	✓		✓****	
S08	✓	✓	✓				✓
S09	✓	✓	✓	✓			✓
S10	✓	✓	✓	✓		✓	✓
S11	✓	✓	✓	✓			✓
S12 – Delivery site 1	✓	✓	✓	✓	✓		
S12 – Delivery site 2	✓	✓	✓			✓	✓
S13	✓	✓	✓	✓			✓
S15	✓	✓	✓			✓	✓

\*Third sector role in the provision of usual care included: referral to or information about local third sector organisations, third sector run group interventions (1 site), and memory cafes.  
 \*\* Other usual care provision included services such as: Admiral nurses, home visits, and co-production

group for people with dementia and supporters/carers. \*\*\* Only relevant to one of the four groups provided at this site. \*\*\*\* The additional group was not run at the same time as JtD groups as the usual facilitators were delivering the JtD intervention.

## References

1. Schuurmans H, Steverink N, Frieswijk N, Buunk BP, Slaets JPJ, Lindenberg S. How to Measure Self-Management Abilities in Older People by Self-Report. The Development of the SMAS-30. Qual Life Res. 2005;14:2215–28
2. Lawton MP, Brody EM. Assessment of Older People : Self-Maintaining and Instrumental Activities of Daily Living 1. Gerontologist. 1969;9(3):179–86
3. Woods RT, Orrell M, Bruce E, Edwards RT, Hoare Z, Hounsborne B et al. REMCARE: Pragmatic Multi-Centre Randomised Trial of Reminiscence Groups for People with Dementia and their Family Carers. Effectiveness and Economic Analysis. PloS One. 2016; 11(4).
4. Devlin NJ, Shah KK, Feng Y, Mulhern B, van Hout B. Valuing health-related quality of life: An EQ-5D-5L value set for England. Health Econ. 2018;27(1):7–22.
5. Mulhern B, Rowen D, Brazier J, Smith S, Romeo R, Tait R, et al. Development of DEMQOL-U and DEMQOL-PROXY-U: Generation of preference-based indices from DEMQOL and DEMQOL-PROXY for use in economic evaluation. Health Technol Assess (Rockv). 2013;17(5):1–140.
6. Rubin DB. Multiple Imputation for Nonresponse in Surveys. New York: John Wiley & Sons, Ltd; 2004.

## 2. Statistical Analysis Plan

### Statistical Analysis Plan

Version 2.0



<b>Trial Title</b>	<b>A randomised controlled trial of the clinical and cost-effectiveness of the Journeying through Dementia intervention compared to usual care</b>
<b>Short Title</b>	<b>Journeying Through Dementia (JtD) Trial</b>
<b>Funder and Reference</b>	<b>NIHR HTA 14/140/80</b>
<b>Sponsor</b>	<b>Sheffield Health and Social Care NHS Foundation Trust</b>

#### *Authored by*

\_\_\_\_\_  
*Ellen Lee*  
*Trial Statistician*  
*Sheffield CTRU, University of Sheffield*

\_\_\_\_/\_\_\_\_/\_\_\_\_  
*Date*

#### *Approvers*

\_\_\_\_\_  
*Prof Stephen Walters*  
*Senior Trial Statistician*  
*ScHARR, University of Sheffield*

\_\_\_\_/\_\_\_\_/\_\_\_\_  
*Date*

\_\_\_\_\_  
*Prof Gail Mountain*  
*Chief Investigator*  
*University of Bradford*

\_\_\_\_/\_\_\_\_/\_\_\_\_  
*Date*

\_\_\_\_\_  
*Prof Catherine Hewitt*  
*TSC Chair and Senior Statistician*  
*University of York*

\_\_\_\_/\_\_\_\_/\_\_\_\_  
*Date*

\_\_\_\_\_  
*Emily Robinson*  
*DMEC Statistician*  
*Kings College London*

\_\_\_\_/\_\_\_\_/\_\_\_\_  
*Date*

## Table of Contents

1.	Introduction and Trial Objectives and Design.....	21
1.1	Brief Trial Background and Primary Research Question .....	21
1.2	Trial Design.....	21
1.3	Documents Guiding the SAP .....	21
1.4	Main Aims and Trial Objectives .....	22
2.	Scope of the SAP .....	22
3.	Outcomes Measures and Timing.....	22
3.1	The Internal Pilot Primary Outcomes .....	22
3.2	Primary Outcome .....	23
3.3	Secondary Outcomes for Participants.....	23
3.4	Secondary Outcomes for Participating Supporter .....	23
3.5	Safety and Harms Outcomes.....	24
4.	Sample Size Estimation .....	24
5.	Trial Features to Minimise Bias .....	25
5.1	Randomisation, Allocation Concealment and Blinding of Outcome Assessments .....	25
6.	Data Monitoring and Interim Analyses .....	25
7.	Data Sources and Data Management.....	26
8.	Definition of Analysis Populations.....	26
8.1	Analysis populations for participants.....	26
8.2	Analysis population for participating supporters .....	27
9.	Outline of the Statistical Analysis.....	27
9.1	Reporting Framework of Trial Data.....	27
9.2	Internal Pilot Analysis.....	27
9.3	The CONSORT Flowchart: Data Completeness and Disposition.....	27
9.4	Demographics and Baseline Characteristics of Participants.....	27
9.5	Characteristics of completers and non-completers.....	29
9.6	Compliance to JtD Intervention.....	29
9.7	Unblinding of Treatment Allocation .....	30
9.8	Efficacy analysis of the intervention.....	30
9.8.1	Primary Endpoint Analysis.....	30
9.8.2	The inclusion of couples into the study .....	30

9.8.3	Sensitivity Analyses for the Primary Endpoint .....	31
9.9	Secondary Efficacy Analyses: Participants .....	31
9.10	Secondary Efficacy Analyses: Participating Supporter .....	31
9.11	Subgroup analysis.....	32
9.12	Assessment of follow up time in relation to outcome .....	32
9.13	Analysis of Safety and Harms.....	33
10.	Detailed Statistical Methods and Calculations.....	33
10.1	Stratification Factors and Multicentre Data .....	33
10.2	Complier Average Causal Effect Inference .....	34
10.3	Mixed effects model checks .....	34
10.3.1	Clustering Implications and Degrees of Freedoms.....	34
10.4	Imputation of missing primary outcome data.....	35
10.4.1	Regression Imputation .....	35
10.4.2	Multiple Imputation .....	35
11.	Data manipulation and definitions .....	35
11.1	Definitions .....	35
11.2	DEMQOL Scoring System.....	36
11.3	PHQ-9 Scoring System .....	36
11.4	GAD-7 Scoring System .....	36
11.5	SMAS Scoring System .....	36
11.6	GSE Scale Scoring System.....	37
11.7	Diener's Flourishing Scale Scoring System.....	37
11.8	IADL Scoring System .....	37
11.9	MMSE Scoring System .....	37
11.10	EQ-5D-5L Scoring system.....	37
11.11	SCQ scoring system .....	37
12.	Implementation of the SAP .....	37
13.	Questionnaire summary table.....	39
14.	Example Tables and Figures.....	40
15.	References .....	50

## SAP amendments since version 1

Version	Date approved	Modifications	Prior to/after blind review	Prior to/after unblind review
2.0		<ul style="list-style-type: none"> <li>- Changed to reflect SAEs not collected for participating supporters (section 9.13)</li> <li>- Changed to reflect DEMQOL and EQ-5D-5L collected either by phone or face to face (sections 11.2, 11.10)</li> <li>- Removed smoking status from example table as not collected (section 14)</li> </ul>	Blind review has not been conducted, however blinded data has been seen in order to start developing programs	Prior to unblind review

## List of Abbreviations

AE	Adverse Event
CACE	Complier Average Causal Effects Analysis
CI	Confidence Interval
CONSORT	CONsolidated Standards Of Reporting Trials
CRF	Case Report Form
CTRU	Clinical Trials Research Unit
DMP	Data Management Plan
eCRF	Electronic Case Report Form
DEMQOL	Dementia related quality of life measure
DMEC	Data Monitoring and Ethics Committee
EQ-5D-5L	Measure of Health Status
GAD-7	Measure of Generalised Anxiety Disorder
GLM	Generalised Linear Model
GSE	General Self-Efficacy Scale
HTA	Health Technology Assessment
HRQoL	Health related quality of life
IADL	Instrumental Activities of Daily Living
ICC	Intraclass correlation coefficient
ICH	International Conference of Harmonisation
IQR	Interquartile range
ITT	Intention-to-treat
JtD	Journeying through Dementia
LTFU	Lost to follow-up
MI	Multiple Imputation
MMSE	Mini Mental State Examination Score
NIHR	National Institute for Health Research
NHS	National Health Service
OR	Odds ratio
PHQ-9	Patient Health Questionnaire
PP	Per-protocol
PROM	Patient reported Outcome Measure
QoL	Quality of Life
RCT	Randomised Controlled Trial
SAE	Serious Adverse Event
SAP	Statistical Analysis Plan
SCQ	Sense of Competence Questionnaire
SD	standard deviation
SMAS	Self-management ability scale
TAU	Treatment As Usual
TMG	Trial Management Group
TSC	Trial Steering Committee

## List of Tables

Table 1: Definitions of Analysis Populations for participants.....	26
Table 2: Baseline Data to be presented.....	28
Table 3: Descriptive JtD intervention compliance summaries.....	29
Table 4: Socio-demographic and characteristics of participants. ....	40
Table 5: Baseline medical history of participants.....	41
Table 6: Baseline characteristics and quality of life of participants for the ITT set. ....	41
Table 7: Baseline characteristics of participating supporters. ....	43
Table 8: Exploratory effect of JtD intervention by subgroups: DEMQOL at 8 months.....	44
Table 9: Continuous baseline characteristics by group and missing data status.....	45
Table 10: Categorical baseline characteristics by intervention group and missing data status. ....	46
Table 11: Primary effectiveness analysis: DEMQOL total score at 8 months.....	47
Table 12: Sensitivity analyses on the primary outcome: DEMQOL total score at 8 months.....	47
Table 13: Exploratory effect of JtD intervention by subgroups: DEMQOL at 8 months.....	48
Table 14: Effectiveness analysis: secondary outcomes at 8 and 12 months.....	48
Table 15: Serious Adverse Events by treatment group .....	49
Table 16: Harms by treatment group: GAD-7 and PHQ-9 .....	49

## List of Figures

Figure 1 CONSORT flow diagram of participants through the trial.....	40
Figure 2 Forest plot of sensitivity analysis on primary outcome: DEMQOL at 8 months.....	47

## Definition of Terms

**Participant-** This refers to a person with dementia who is participating in the trial.

**Participating supporter-** This is a family member, friend or neighbour that provides support to a person with dementia. They may be known as a ‘carer’. In the trial, participating supporters are people that have consented into the trial to complete outcome measures. They may also help a person with dementia participate in the trial, such as liaising with researchers to organise visits; and participating in the intervention if allocated to receive it.

**Supporter-** This is a family member, friend or neighbour that provides support to a person with dementia. They may be known as a ‘carer’. In the trial, supporters are people that may be helping a person with dementia participate in the trial, such as liaising with researchers to organise visits or attending the intervention if allocated to receive it. However supporters are not participating in the trial themselves, for example no outcome data is collected from them.

## 1. Introduction and Trial Objectives and Design

This Statistical Analysis Plan (SAP) provides detailed guidance of the statistical analysis for the Journeying through Dementia (JtD) trial. This section provides a brief background of the trial, the primary research question under investigation, study design used to address the research questions, and key documents guiding the development of this SAP.

## 2. Brief Trial Background and Primary Research Question

There is an increasing burden of dementia in the UK and globally impacting those living with the condition, and their family carers, services and the economy. Whilst existing research has provided insights into the potential benefits of promoting self-management for people with dementia, there is an absence of robust evidence through full-scale RCTs. This means that it is difficult to establish the effectiveness and cost-effectiveness of such interventions, particularly in comparison to usual care.

The JtD intervention has been designed to support people in the early stages of dementia and improve their quality of life (QoL) by promoting self-efficacy and assisting them to continue to participate in life and maintain their independence. The JtD involves individuals in the early stages of dementia participating in 12 facilitated weekly group sessions and in 4 individual sessions with one of the facilitators. It is expected that the timing of the individual sessions will be one before, one after and two during the course of the group sessions. The group is encouraged to select the content of their sessions from a range of topics including strategies to manage memory challenges, engaging in hobbies/interests and ways of maintaining physical and mental wellbeing. An essential component is the enactment of activities in the community with support from each other. During individual sessions people with dementia are assisted to work on individual needs and goals. Participants are not necessarily required to nominate a supporter (family member or friend who provides them with support) to take part, but if supporters are involved they are invited to join group sessions one, six and twelve and can participate in the individual sessions with the person with dementia if agreed.

This trial therefore aims to determine the clinical and cost-effectiveness of the JtD intervention for people in the early stages of dementia. The trial is funded by the National Institute for Health Research (NIHR) Health Technology Assessment (HTA) Programme, and the Sheffield Health and Social Care NHS Foundation Trust acts as the trial sponsor.

## 3. Trial Design

This study is a pragmatic, two-arm, parallel group, superiority, individually and equally randomised controlled trial (RCT) comparing the JtD intervention with usual care to determine benefit for people in the early stages of dementia. Although the study is an individually RCT, the delivery of the JtD intervention in one arm of the trial is group based. That is, groups of 8 to 12 participants randomised to the JtD intervention will attend 12 weekly facilitated meetings in local venues in order to receive the intervention. The trial has been designed with an internal pilot phase during the first 8 months of active recruitment which will assess the feasibility aspects of conducting the main trial as guided by pre-planned STOP/GO criteria.

For consistency throughout this SAP, the active intervention group shall be referred to as “JtD” and the control group as “Usual Care”.

## 4. Documents Guiding the SAP

This SAP is written in conjunction with the International Conference of Harmonisation (ICH) topic E9 on Statistical Principles for Clinical Trials (ICH, 1998), applicable Standard Operating Procedures from the Sheffield Clinical Trials Research Unit (CTRU), (particularly ST001), and the Trial Protocol and related amendments.

## 5. Main Aims and Trial Objectives

As highlighted in Section 2, the primary aim of the trial is to determine the clinical and cost effectiveness of the JtD intervention compared to usual care for people in early stages of dementia. The specific objectives are to:

- a) conduct an internal pilot RCT to assess the feasibility of rates of recruitment considered,
- b) proceed to the full trial if the internal pilot trial feasibility criteria are met,
- c) conduct fidelity checks regarding the delivery of the JtD intervention,
- d) conduct an embedded qualitative nested study to explore issues concerned with intervention delivery,
- e) identify how the intervention might be realistically delivered through services.

## 6. Scope of the SAP

This trial has been designed with an internal pilot phase to assess the feasibility of patient recruitment. Furthermore, there is a fidelity study to assess the delivery of the JtD intervention and an embedded qualitative sub-study to explore issues concerned with the delivery of the intervention. The trial is also designed with health economic evaluation to address the cost effectiveness of the JtD intervention. This SAP, however, focuses only on addressing the internal pilot objectives and clinical effectiveness research questions of the trial. Therefore, the fidelity and embedded qualitative studies and health economic evaluation aspects which are out of scope of this SAP will be addressed elsewhere.

## 7. Outcomes Measures and Timing

This section describes the outcome measures, which are used to evaluate the feasibility, primary, and secondary objectives of the trial relating to the clinical effectiveness of the JtD intervention compared to usual care. The timing of all outcome measures is explicitly stated. Outcomes relating to fidelity and embedded qualitative studies are excluded, as they are out of scope of this SAP. Although health economic evaluation is out of scope of this SAP, related outcomes are stated here for completeness as their analysis will be reported elsewhere.

### 8. The Internal Pilot Primary Outcomes

The following feasibility outcomes will be reviewed after 8 months of active recruitment to assess the STOP/GO criteria:

- Recruitment of a minimum of 113 participants across the six pilot sites by the end of the fifth month of active recruitment (75% of the 150 target).
- Recruitment of a minimum of 12 facilitators (two facilitators identified at each of the six pilot sites by the start of active recruitment to deliver the intervention).

- No more than two of the six planned groups in the internal pilot with less than four participants registered for the group by the sixth month of active recruitment.

## 9. Primary Outcome

To address the primary clinical effectiveness research question of the JtD intervention, the primary endpoint assesses the health related QoL in participants measured by the DEMQOL (Smith et al, 2005) which is evaluated at 8 months post randomisation. The scoring system to generate summary scores for analysis is detailed in Section 46.

## 10. Secondary Outcomes for Participants

Seven secondary PROMs are assessed at 8 months post-randomisation focusing on participants. Scoring systems used to compute summary measures and dealing with missing data are detailed in Section 46:

- PHQ-9; measure depression severity (Hancock et al, 2009)
- Generalised Anxiety Disorder Assessment (GAD-7); severity measure of generalised anxiety disorder (Wild et al, 2014)
- Instrumental Activities of Daily Living (IADL); measures independent living skills (Gold, 2012)
- General Self Efficacy (GSE); measures general perceived self-efficacy (Schwarzer et al, 1995)
- Diener's Flourishing Scale; measures psychological flourishing (Diener et al, 2010)
- Self-Management Ability Scale (SMAS); measures self-management abilities (SMA) (Schuurmans, 2005)
- EQ-5D-5L; measures health outcome (Brazier et al, 2007)

Additional secondary PROMs and resource use outcomes are assessed at 12 months post-randomisation to further explore the clinical and cost-effectiveness research questions.

- DEMQOL
- EQ-5D-5L
- Health and social care resource use questionnaire (for use in health economic evaluation)

## 11. Secondary Outcomes for Participating Supporter

The following secondary outcomes will be assessed at 8 months post randomisation focusing on the participating supporter:

- PHQ-9

- b) Sense of Competence Questionnaire (SCQ); measures sense of competence in caregivers  
(Vernooij-Dassen, 2003)
- c) EQ-5D-5L

## 12. Safety and Harms Outcomes

There are few anticipated adverse effects of the JtD intervention. Adverse Events (AEs) are not anticipated as a consequence of the intervention and thus will not be recorded. Serious Adverse Events (SAEs) will be recorded for all participants. Six categories of SAEs will be recorded during follow up:

- (a) results in death
- (b) is life-threatening (subject at immediate risk of death)
- (c) requires hospitalisation or prolongation of existing hospitalisation
- (d) results in persistent or significant disability or incapacity
- (e) consists of a congenital anomaly or birth defect
- (f) is otherwise considered medically significant by the investigator

## 13. Sample Size Estimation

As stated in Section 9, the primary outcome for the study is the mean DEMQOL score 8 months post randomisation. If we assume a standard deviation of 11 points for the DEMQOL, and that a mean difference of 4 or more points is clinically and practically important (Mulhern, 2013). The sample size has been calculated to have a 90% power of detecting this 4-point difference or more, if it truly exists, which is equivalent to a standardised effect size of 0.36 in group mean scores at 8 months, as being statistically significant at 5% two sided level. As the JtD intervention is a facilitator led group intervention, the success of the intervention may depend on the facilitator delivering it so the outcomes of the participants in the same group with the same facilitators may be clustered. For an individually RCT without adjustment for clustering by facilitator, the target sample size would be 160 per arm (a total of 320). Assuming an average cluster size of 8 dementia patients per facilitated group (Mountain, 2017) and an intra-cluster correlation (ICC) of 0.03, this will inflate the sample size by a design effect of 1.21; to 194 per group (a total of 388) with valid primary outcome data. We further assumed 20% lost to follow-up. Therefore, given these calculations, the trial target sample size is to randomise to 243 participants in each arm (a total of 486).

The Dementia Quality of Life measure (DEMQOL) (Mulhern et al 2013) is a 28 item self-completed quality of life questionnaire for people with dementia. The 28 items are scored on a 1 to 4 Likert scale (a lot; quite a bit; a little; not at all) and the total score ranges from 28 to 112. There is no reported or established minimum clinically important difference (MCID) for the DEMQoL so it is difficult to specify a target difference that is clinically and practically important. Data from the development and validation of the preference/utility based DEMQOL-U outcome (Mulhern et al 2013) ) which compared the performance of the DEMQOL (and DEMQOL-U) with other patient reported outcome measures for people with dementia; for example MMSE (Mini Mental State Examination) (Folstein et al 1975); BADLS (Bristol Activity of Daily Living Scale) (Bucks et al 1996), NPI (Neuropsychiatric Inventory) (Cummings et al 1994) with established MCIDs reported changes on the DEMQOL of 5.4; 7.5 and 6.4 points respectively for patients who improved by more than the MCID on the anchor measure. This suggests that our proposed 4-point difference, although small is likely to be of clinical and practical importance.

## 14. Trial Features to Minimise Bias

This section describes design measures put in place to avoid the potential of bias in assessing the effectiveness of the JtD intervention, focusing on randomisation, its concealment, and blinding of outcome assessors and other research staff. Additional measures adopted to minimise bias during statistical analyses, such as dealing with missing data, compliance to the JtD intervention, clustering in one arm and stratification during randomisation are addressed in Section 38.

## 15. Randomisation, Allocation Concealment and Blinding of Outcome Assessments

Participants are randomised 1:1 to receive either JtD intervention or Usual Care. Randomisation stratified by stratification site (see section 46 for a definition of stratification site) and blocked by a fixed block size will be undertaken using a web-based randomisation system. The objective is to minimise allocation predictability and imbalance in the distribution of participants between intervention arms and also within delivery site. A fixed block size was used to ensure the imbalance between treatment groups was controlled and minimal; at the start of the study only two JtD intervention groups per site were deemed feasible, and a maximum number of 13 participants allowed in each JtD intervention group. The risk of predicting the allocation ratio in this study is minimal as randomisation is done centrally by Sheffield CTRU staff members and research assistants responsible for recruiting participants at site are blinded to treatment allocation.

Due to the group nature of the intervention, a process of delayed randomisation will be employed whereby participants are not randomised at the point of consent but after the collection of baseline measures (less than 2 months before the intervention wave at that site begins). This will enable operational feasibility in the organisation of the JtD group sessions to minimise delay in intervention delivery. It also means that the time between randomisation and starting treatment is roughly similar (within 8 weeks) across participants, and that treatment occurs close to the beginning of the baseline to 8 month follow up window. In the event of a couple in the same household both consenting to take part in the study the pair will be randomised as a couple and not separately i.e. to both get the intervention or to both get usual care.

A number of measures have been undertaken to minimise potential bias. Randomisation and its concealment will be done using a centralised web-based randomisation system. Furthermore, since the trial uses a PROM as a primary outcome, it is imperative to maintain the blinding of outcome assessors as much as possible during the course of the trial. In this regard, randomisation will be performed centrally by staff at the Sheffield CTRU rather than site staff. Furthermore, to ensure the outcome assessors are blinded to group allocation, the trial support officer, clinical research assistant, trial manager or other member of the study team who will not be conducting outcome assessment will enter the participants' details onto a remote web-based randomisation system. Trial Statisticians and Health Economists will be blinded during the course of the trial until data freeze. Any 'unblinding' cases for various circumstances will be reported as described in Section 28 for transparency and to enhance the interpretation of results.

The reports to the Data Monitoring and Ethics Committee (DMEC) will be prepared by Data Management who are unblind. Unblinded statistical reports may also be provided to the DMEC on their request, as guided by the DMEC Charter by a Sheffield CTRU Statistician who is external to the trial.

## 16. Data Monitoring and Interim Analyses

The conduct of this trial will be guided by three committees as governed by Sheffield CTRU SOPs, trial protocol, and the DMEC Charter. The committees are:

- a. Trial Management Group (TMG)
- b. Trial Steering Committee (TSC)
- c. DMEC

This trial has been designed as a fixed sample size design with one formal statistical analysis at the planned scheduled end. Therefore, there are no planned interim analyses to allow early stopping using formal statistical rules. However, the trial will be independently monitored by the DMEC within the premise of the related Charter agreed and signed by all the members. A recommendation to stop the trial could be made by the DMEC based on safety reasons as stipulated in the DMEC Charter. Although unlikely to happen, on DMEC's discretion, they may request futility analysis using stochastic curtailment but there is no option for efficacy early stopping.

## 17. Data Sources and Data Management

Data used in this study will come from data entered onto Case Report Forms (CRFs) and questionnaires. The data are stored on the Sheffield CTRU database system, with the exception of the randomisation list which is held on the CTRU's randomisation system. The CTRU database system has restricted access to certain trial staff depending on their trial related duties and responsibilities. The data management unit of the Sheffield CTRU will validate and query electronic data for inconsistencies during the course of the trial as governed by the processes and procedures stipulated in the Central Data Validation SOP (DM005). The Trial Statistician will conduct any additional validation checks where appropriate before the data lock and sign off guided by the relevant SOPs (DM005, DM012).

## 18. Definition of Analysis Populations

This section states and defines the primary analysis populations and other secondary analysis populations which will be used mainly for sensitivity analysis.

## 19. Analysis populations for participants

To minimise potential selection bias and provide unbiased estimate of the treatment effect and to preserve randomisation; primary analysis population will be based on Intention-to-treat (ITT) principle as defined in Table 1. In order to truly implement ITT analysis, missing primary outcome data will be imputed and included in a sensitivity analysis, as described in Section 31.

Table 1: Definitions of Analysis Populations for participants.

Analysis Population	Patient inclusion criteria
Intention-to-treat (ITT)	<p>(1) All participants randomised to receive either JtD intervention or Usual Care <b>and</b>,</p> <p>(2) Consented to take part in the study <b>and</b>,</p> <p>(3) Treatment assignment during analyses is as allocated at randomisation regardless of what happens after randomisation (such as protocol deviations and withdrawals) <b>and</b>,</p> <p>(4) With complete DEMQOL data at 8 months post randomisation (see Section 48 for a definition of a complete score).</p> <p>Excluding participants who withdraw before randomisation, and including participants found to be ineligible post randomisation (Altman, 1991).</p>
CACE	a) A subset of the ITT analysis population

- 
- b) Subgroup of participants who are believed to comply with the JtD intervention as described in Section 27 and 40,
  - c) Excludes ineligible participants randomised by error
  - d) Includes participants who were randomised to usual care but received and complied with JtD intervention
- 

## 20. Analysis population for participating supporters

This includes participating supporters who met eligibility criteria as stated in the protocol and provided informed consent. Since this is a supplementary analysis, only participating supporters with available follow-up data will be analysed; thus, no imputation will be employed.

## 21. Outline of the Statistical Analysis

This section outlines the statistical analysis framework to be adopted beginning with how the trial data and results will be reported. The description of the statistical methods used to analyse outcomes to address the research questions are given in order of importance, starting with the primary outcome followed by secondary outcomes, including safety. Dummy tables can be found in section 60 and are provided as an example and to guide the Trial Statistician(s) during analysis and reporting.

## 22. Reporting Framework of Trial Data

The analysis of trial data and reporting will be guided by the revised CONSORT statement for individually randomised parallel group trials and the extension to the consort statement for non-pharmacologic treatments (Boutron et al, 2008).

It should be highlighted that the trial is also partially clustered in one arm. Thus summaries relating to this clustering within the JtD arm will be reported; for example, number of JtD courses and their attendance, and the observed ICC (with 95% CI) to help the planning of future related partially clustered trials.

## 23. Internal Pilot Analysis

The Trial Steering Committee (TSC), tasked to 'independently' oversee conduct of the trial on behalf of the funder and sponsor, will assess the feasibility outcomes at the end of the pilot phase. They will assess whether the trial should continue in light of the feasibility results against the STOP/GO criteria. The results and recommendations will be communicated to the funder (NIHR HTA).

## 24. The CONSORT Flowchart: Data Completeness and Disposition

The reporting of data completeness is an integral part of good practice in trial reporting. Guided by the CONSORT statement for individually randomised parallel group RCTs, a flowchart showing participants (not the participating supporters) from screening, during follow-up, and to inclusion in the primary analysis will be constructed as illustrated in Figure 1. This will also be made available to the trial monitoring committees during the trial. Summaries may be presented by centre and/or overall (and treatment group where appropriate). Data completeness will be based on the primary outcome, DEMQOL score.

## 25. Demographics and Baseline Characteristics of Participants

Summaries of the baseline variables relating to socio-demographics, medical history and characteristics of participants as captured on the CRFs will be presented by intervention group and overall (as in Table 4, Table 5, Table 6, Table 7). The list of baseline variables to be presented is given in Table 2.

No statistical significance testing will be done to test baseline imbalances between the intervention arms but any noted differences will be descriptively reported. Baseline imbalances will be considered for adjusted analysis as part of the sensitivity analysis described in Section 29; (the TMG will discuss baseline

imbalances, and deemed potentially clinically relevant by the chief investigator will be included in the sensitivity analysis).

Summaries will be presented depending on the distribution of the variable under consideration. For example, minimum, maximum, mean and standard deviation (SD) or median, Interquartile Range (IQR), minimum and maximum will be presented for continuous variables depending on the skewness of the underlying data. Categorical variables will be summarised using numbers and related percentages in each category by intervention group and overall. As for count variables, a decision on the presentation approach will be made based on the underlying distribution of the pooled data. For instance, if the maximum number of counts is small, then a categorical variable will be derived and presented accordingly. Otherwise, the median (IQR) of the distribution of the count variable will be presented.

*Table 2: Baseline Data to be presented*

Participant Socio-demographics	Categorical variables <ul style="list-style-type: none"> <li>• Delivery site</li> <li>• Sex</li> <li>• Ethnicity</li> <li>• Living with others (yes/no)</li> <li>• Lives with (spouse, children, etc)</li> <li>• Accommodation type</li> </ul> Continuous variables <ul style="list-style-type: none"> <li>• Age (years)</li> </ul>
Participant medical history	Categorical variables <ul style="list-style-type: none"> <li>• Specific health problem</li> <li>• Type of dementia diagnosis</li> </ul> Continuous variables <ul style="list-style-type: none"> <li>• Length of time since dementia diagnosis (years)</li> </ul>
Participating Supporter baseline characteristics	Categorical variables <ul style="list-style-type: none"> <li>• Sex</li> <li>• Relationship to person with dementia</li> <li>• Length of time caring (categories)</li> <li>• Living with person with dementia</li> </ul>
Participant Baseline quality of life	Continuous outcomes <ul style="list-style-type: none"> <li>• DEMQOL (total score and Q29)</li> <li>• EQ-5D-5L (value index and VAS)</li> <li>• PHQ-9</li> <li>• GAD-7</li> <li>• GSE</li> <li>• Diener's Flourishing Scale</li> <li>• SMAS</li> <li>• IALD</li> </ul>
Participating Supporter baseline quality of life	Continuous outcomes <ul style="list-style-type: none"> <li>• PHQ-9</li> <li>• EQ-5D-5L (value index and VAS)</li> <li>• SCQ</li> </ul>

## 26. Characteristics of completers and non-completers

The objective of this section is to explore the pattern of missing data of participants and whether completers are systematically different from non-completers. Completers are participants meeting the ITT criteria, who have complete primary endpoint data (DEMQOL at 8 months from randomisation).

Descriptive statistics of baseline variables will be presented stratified by the intervention group and missing data status as illustrated in *Table 9* and *Table 10*. The baseline variables include age, sex, smoking status, type of dementia and QoL measures. Other variables presented in baseline tables may be considered at the discretion of the Trial Statistician. The relationship between treatment compliance and missing data status will be tabulated for the participants in the Intervention group. Predictors of missing data will be included in multiple imputation equations.

## 27. Compliance to JtD Intervention

This section aims to summarise the attendance of participants to group and individual sessions which are part of the JtD intervention. In addition, details about the accompanying persons and their relationship with the participants will also be summarised. Participants in the JtD arm are expected to attend 12 group sessions and 4 individual sessions; a total of 16 possible sessions. Participants are able to invite a supporter to participate in group sessions 1, 6 and 12, and in the individual sessions if the participant finds this helpful in achieving their goals. As stated in the protocol, a person with dementia will be classified as having received the JtD intervention if they attended at least 10 of the possible 16 sessions (62.5%). The list of JtD intervention compliance summaries is presented in *Table 3*.

*Table 3: Descriptive JtD intervention compliance summaries*

General summary (N)	<ul style="list-style-type: none"> <li>Total number of JtD groups</li> </ul>
Continuous (mean, SD, median, IQR, min, max)	<ul style="list-style-type: none"> <li>Session size</li> <li>Number of sessions attended (group + individual)</li> <li>Number of group sessions attended</li> <li>Number of individual sessions attended</li> <li>Proportion of group sessions accompanied</li> </ul>
Categorical (n, %)	<ul style="list-style-type: none"> <li>Reason for non attendance</li> <li>Reason for intervention withdrawal</li> <li>Number of group sessions accompanied (0,1,2,3)</li> <li>Number of individual sessions accompanied (0,1,2,3,4)</li> <li>Accompanying person relationship to participant (participating supporter, non-participating supporter, consultee or other)</li> </ul>

Data visualisation approaches will be used to present data in appropriate figures or graphs. Graphical methods and correlation will be used to investigate a relationship between the number of JtD sessions attended and DEMQOL score at 8 months.

Exploratory descriptive analysis will compare those that ‘received’ the intervention to those that did not (within the intervention arm) with respect to baseline data (age, type of dementia etc.) and socio-economic status as defined by the ONS Index of Multiple Deprivation calculated using participant’s postcode data.

## 28. Unblinding of Treatment Allocation

As highlighted in Section 15, outcome assessors will be blinded to intervention allocation in order to minimise operational bias. For some reasons, unblinding of treatment allocation may happen. It is therefore important to report the following summaries by intervention group:

- a) Number and proportion of unblinded cases,
- b) Who was unblinded (outcome assessors, Trial Statisticians, TSC, Health Economist, etc),
- c) Method of unblinding (face to face, email, by observation, etc),
- d) Source of unblinding (facilitator, participant, DMEC, etc) ,
- e) Reason for unblinding,
- f) Suspected or real (see section 47).

## 29. Efficacy analysis of the intervention

This section describes the statistical methods to be used for the primary analysis and related sensitivity analyses to address the primary research question. The Usual Care arm will be the reference group for the analysis unless stated otherwise.

## 30. Primary Endpoint Analysis

The primary aim of the trial is to determine the benefit of an occupational therapy based self-management intervention (JtD) compared to Usual Care for people in the early stages of dementia. To address this aim, the primary analysis will compare mean patient reported DEMQOL total scores at 8 months post randomisation between the JtD intervention and Usual Care groups using a mixed effects linear regression model adjusted for DEMQOL baseline total score and stratification site, and allowing for the clustering of the outcome by the JtD intervention group (Baldwin et al., 2011; Roberts et al., 2015; Walters, 2010). The trial design is partially nested where there is clustering in one arm (JtD) only. Each participant in the unclustered Usual Care arm will be treated as a cluster of size one (a singleton). The cluster indicator will be treated as a random effect and stratification site (a stratification and centre effect variable used for randomisation) as a fixed factor. These two factors will be included in all statistical models.

A partially clustered mixed effects linear regression model with homoscedastic errors as well as a heteroscedasticity mixed effects linear regression model will also be considered to account for potential differential variability of outcomes between the two intervention groups. The final model will be selected based on whether the assumption of homoscedastic errors is reasonable. Degrees of freedom will be computed using Satterthwaite approximation to account for a potential of a small to moderate number of clusters (Satterthwaite, 1946). A 95% confidence interval (CI) for the mean difference in DEMQOL total scores between the JtD and Usual Care arm will be reported together with the associated P-value.

A further adjusted analysis may also be performed alongside this baseline DEMQOL and site adjusted analysis depending on the observed degree of imbalance of important baseline covariates which are or may be of potential prognostic importance using a mixed effects linear regression model as described for the primary analysis. Additional baseline covariates of prognostic importance will include age, sex, PHQ-9 (total score), and GAD-7 (total score).

## 31. *The inclusion of couples into the study*

It is possible that two people living with dementia in a couple within the same household will want to take part in the study. These participants will be randomised as a couple (i.e. both to intervention or both to control). In the event that there are more than 10 couples (20 participants) the primary and secondary

analyses will be changed to take into account the hierarchical or clustered nature of the data. A multi level mixed effects model will be used; the random effects will be JtD intervention groups (top level) and couple/singles (lower level). Individual participants who are not part of a couple will be treated as clusters of size one. We found in a previous similar study (18 couples recruited, total sample size 288) that there was no difference in the primary effectiveness analysis allowing for clustering by couples or not (Mountain, 2017).

If there are 10 couples or fewer, the primary analysis will be as is stated in section 30; a random effects model with one level of clustering (the JtD intervention groups). This means the clustering of outcomes within couples will be ignored. A sensitivity analysis will be performed using the multi level hierarchical model that includes couple/singles (as clusters of size 2 or 1) will be performed to check the similarity of results.

### 32. *Sensitivity Analyses for the Primary Endpoint*

Sensitivity analyses on the primary outcome will be undertaken based on imputation, CACE and mistimed measurements (MM) populations. The CACE population has been described in Section 19. Additional imputation and MM populations are defined as follows:

- Imputation: All participants who were consented and randomised (ITT) including those with missing primary outcome data, excluding those that died before 8 month follow up. For participants that do not have complete primary outcome data, their DEMQOL 8 month score will be imputed by regression imputation and multiple imputation (MI). See section 43 for details of these imputation methods.
- Mistimed Measurements (MM); a subset of the ITT population which excludes DEMQOL measurements which were collected outside the acceptable window. Ideally, outcomes will be collected < 2 weeks pre- and <8 weeks post-date they are due. Outcomes collected outside this window will be excluded for this sensitivity analysis.

The primary analysis will be repeated for these data sets and displayed alongside the intention to treat analysis results. Details of CACE analysis are provided in Section 40. Sensitivity analyses results will be reported as shown in Table 12 (including the primary ITT set) and presented using a forest plot (Cuzick, 2005), as illustrated in Figure 2.

### 33. *Secondary Efficacy Analyses: Participants*

Secondary outcomes including the EQ-5D-5L, PHQ-9, GAD7, GSE, Diener's Flourishing Scale, SMA and IADL at 8 months post-randomisation will be compared between the intervention and control groups using a mixed effects linear regression model as for the primary outcome. A 95% CI for the mean difference in this parameter between the treatment groups will also be calculated together with the associated P-value, and 95% CI. A similar approach will be used to compare secondary outcomes at 12 months post randomisation.

### 34. *Secondary Efficacy Analyses: Participating Supporter*

Outcome measures for the participating supporters will be collected at baseline and 8 months. This includes the PHQ-9, EQ-5D-5L and SCQ. The 8 month outcomes will be compared between the intervention

and control groups using a least squares linear regression model. Participating supporters are invited to attend up to 3 sessions, correlation between participating supporters within the same cluster is likely to be minimal, hence a simpler model is preferred and no adjustment for JtD cluster group will be made. The mean difference in outcome with associated 95% CI and P-value will be presented for: a) analysis adjusting for the baseline outcome (i.e. PHQ-9) and stratification site and b) adjusted analysis with additional covariates (in addition to a)) if baseline imbalances have been observed.

### 35. Subgroup analysis

The main objective of this section is to explore heterogeneity in the intervention effects in pre-specified subgroups. An exploratory subgroup analysis will be performed using mixed effects linear regression with the primary outcome, the mean DEMQOL score at 8-month post-randomisation, as the response. An interaction statistical test between the randomised intervention group and subgroup will be used to directly examine the strength of evidence for the difference between treatment group (JtD versus Usual care) varying between subgroups. Two subgroups of interests for exploratory analysis have been pre-specified based on clinical advice:

- Type of dementia
- Presence of participating supporter (yes/no, see section 47).

Subgroup analysis will be performed regardless of the statistical significance on the overall intervention effect. The mean difference in DEMQOL (95% CI) will be computed for each subgroup category and visually displayed using a forest plot (Cuzick, 2005). The regression coefficient for the interaction between treatment group and subgroup will be presented with the associated confidence interval and P-value. We will not calculate separate p-values within each subgroup category (Assmann et al., 2000; Pocock et al., 2002; Wang et al., 2007). Results will also be presented as shown in Table 8. No adjustment for multiple testing will be made to the type 1 error rate.

The categorisation into subgroups for dementia type is not straightforward; dementia diagnosis depends on the methods used at each site and we anticipate that a large amount of participants will have Alzheimer's, which may leave the other subgroups small. Four options are recorded in the CRF (Alzheimer's, Vascular dementia, Mixed Alzheimer's/vascular dementia, other (with free text details)). Subject to adequate group sizes, three categories of dementia will be used for the formal subgroup analysis (Alzheimer's VS any vascular dementia (either vascular dementia or mixed dementia) VS other). A sensitivity analysis using different clinically relevant group combinations will be performed (using the groups; Alzheimer's VS vascular dementia VS mixed + other, and; Alzheimer's + mixed VS vascular VS other).

Descriptive statistics of Mean DEMQOL per group and treatment difference with 95% confidence interval will be presented for each dementia type presented including smaller groups created based on coding the 'other' category (such as dementia with Lewy bodies).

### 36. Assessment of follow up time in relation to outcome

A process of delayed baseline and randomisation will be implemented on this study to ensure that the delay between baseline measures and a JtD group starting is within 2 months. This means the time between conducting baseline measures and a course starting will vary across participants. There are also possible delays between baseline and randomisation as randomisation is performed centrally at Sheffield. The time between baseline, randomisation and starting treatment (for JtD arm) will be summarised by arm. Two sensitivity analyses will be conducted on the primary outcome:

- A covariate of days between baseline and group starting (which will be 0 for the control group) will be included in the analysis model.
- A covariate of days between baseline and 8 month post randomisation follow up visit will be included in the analysis model.

### 37. Analysis of Safety and Harms

The following descriptive statistics of serious adverse events (SAEs) will be calculated and reported by treatment group and overall. SAEs will be reported on an intention to treat basis (i.e. according to the group to which the participant was randomised) for participants only. No formal statistical tests will be performed on SAE data. All detailed listings of SAEs will be made available to the DMEC for each meeting and on their request at any point during the trial. All SAEs will be reported on an intention to treat bases (i.e. according to the group to which the participant was randomised).

The Following summaries will be presented (overall and by group)

- Number (%) participants experiencing  $\geq 1$  SAE
- Number of all SAEs (including repeated events)
- Seriousness (Death, Life threatening, Inpatient hospitalisation, Prolongs hospitalisation, Persistent or significant disability/incapacity, congenital abnormality/birth defect)
- Intensity (mild, moderate, severe)
- Relationship to study intervention (definite, probable, possible, unlikely, unrelated, not assessable)

In order to further investigate potential harms, the proportion of participants and participating supporters considered to be suffering from moderately severe depression symptoms (based on a PHQ-9 score of 15 or above; Kroenke, 2001) will be presented by treatment group and overall for baseline, 8 and 12 months. The proportion of participants considered to be suffering from moderate/severe levels of anxiety (based on GAD-7 score of 10 or above; Spitzer, 2006) will be presented in the same manner.

### 38. Detailed Statistical Methods and Calculations

Due to the nature of the study design, there are a number of statistical issues that should be addressed or accounted for during statistical analyses. This section aims to highlight these issues and recommend a framework to address them.

#### 39. Stratification Factors and Multicentre Data

The stratification site within recruiting trusts is the only stratification variables. Few recruiting trusts are expected to have more than one stratification site. Adjusting for stratification site during the analysis will account for stratification and also potential centre effect while preserving degrees of freedoms.

#### 40. Complier Average Causal Effect Inference

One research question of interest is on the nature of the relationship between adherence to the JtD intervention and benefits of dementia self-management as assessed by health related QoL (DEMQOL). CACE analysis is an attempt to compare the protocol compliers in the JtD group (those that attended at least 10 out of the possible 16 sessions, as detailed in Section 27) to those in the usual care group who are 'likely' to have complied had they been randomised to JtD intervention. CACE analysis will be performed in the following steps (Peng et al, 2004):

1. Using participants in the JtD group, derive a logistic regression model to predict the probability of being a non-complier (i.e. attending less than 10 of the 16 sessions). Possible predictor covariates are age, sex, type of dementia, baseline QoL.
2. Apply these predictions to the Usual Care group, so that each participant is given a probability of receiving the JtD intervention as planned (if they had been randomised to receive it) which is based on their covariates.
3. For each participant in the Usual Care group calculate a re-weighted outcome defined as the original outcome multiplied by the predicted probability of receiving as planned.
4. Compare the subset of participants in the JtD group that are deemed to have complied with intervention (attending at least 10 sessions) with the re-weighted outcomes amongst participants in the Usual Care group.

CACE analysis will be conducted by a two stage regression, the first will use logistic regression (excluding clustering adjustment for JtD group) but including site. This is because the clustering is one-sided (only in the JtD arm) so will not add to the predictive model. The second model used in step 4 will be the mixed effects model as used in the primary analysis.

#### 41. Mixed effects model checks

Model goodness of fit will be investigated via graphical methods (e.g. histograms of residuals and scatterplots of residuals vs. covariates) for primary and secondary endpoints. Checks will be made on whether a homoscedastic errors assumption is reasonable. Influential observations and outliers will also be investigated and sensitivity analyses at the discretion of the trial statistician will be undertaken and reported.

#### 42. Clustering Implications and Degrees of Freedoms

First, although this is an individually RCT, part of the JtD intervention (12 of the 16 sessions) is delivered in groups of ideally 8 to 12 participants. Therefore, outcomes of participants who attend group sessions together are more likely to be similar to each other compared to their counterparts from other group sessions. Furthermore, given that this form of clustering which will happen only in the JtD intervention

arm, there are different ways of coding the unclustered Usual Care arm; treating each participant as a single cluster (singletons), all participants as one cluster, or random creation of artificial clusters (Flight, 2016). Another issue is that the appropriate computation of the degrees of freedoms to preserve the type I error is influenced by the number of clusters (Kahan et al., 2016). In summary, statistical models used for the analyses will account for the clustering in one arm of the trial, employ appropriate coding of the unclustered Usual Care arm, and use appropriate approaches to compute degrees of freedoms as supported by current statistical literature.

#### 43. Imputation of missing primary outcome data

For sensitivity analysis, imputation will be used in order to obtain complete eight month DEMQOL data. No imputation will be performed for participants that have died before the eight month follow up assessment. Missing data will be imputed using at least two methods; regression imputation and multiple imputation.

#### 44. Regression Imputation

Regression imputation will be used to impute missing eight month DEMQOL dimension data. A regression equation that accounts for participant baseline data (age, sex, baseline DEMQOL) will be used to fill in missing values.

#### 45. Multiple Imputation

One hundred multiple imputation data sets will be created using chained equations. The multiple imputation equation will include baseline data (age, sex, baseline DEMQOL), prognostic factors (type of dementia, medical history) and predictors of missing data (see section 26) to make the Missing at Random assumption as plausible as possible. A conservative approach will be adopted and treatment group will be excluded from the imputation model.

#### 46. Data manipulation and definitions

#### 47. Definitions

- For the sake of subgroup analysis presence of participating supporter (yes or no) is defined based on the eligibility assessment question “Is there a participating supporter; yes/no”.
- The unblinding of outcome assessors is defined as real when the outcome assessor has correctly identified the allocation group and suspected when the outcome assessor is believed to be unblinded but failed to identify the group allocation.
- Stratification site: Each site is expected to run (at least) two waves of the JtD intervention, potentially at multiple delivery sites. In cases where participants could attend either delivery sites within a site (for example Sheffield) then the stratification site is identical to the site. In cases where sites are recruiting participants from two distinct delivery sites and the participants are only expected to attend an intervention at that delivery site then each delivery site is defined as a stratification site (for example Teas Esk and Wear (the site) is recruiting at both Harrogate and York the delivery sites). The list of stratification sites will be recorded in the SCRAM randomisation system.

#### 48. DEMQOL Scoring System

DEMQOL may be collected over the telephone or face-to-face, the scoring and analysis will be consistent regardless of the method of collection but we will note the number of instances of telephone collection. A total score based on the first 28 items (excluding the 29<sup>th</sup> item) is calculated by summing the items. The score ranges from 28 – 112 and a higher score represents better health related quality of life. For a dementia patient record to be classified as valid, at least 50% of the first 28 items (at least 14 items) must have responses (Smith et al., 2005). Any records with less than 50% completed items will be deemed invalid and excluded from any analysis. The following recommended approach is adopted to deal with missing item data among those with at least 50% item responses (Bsmsacuk, 2016):

- Calculate the participant specific mean based on items with responses
- Impute the missing items with the patient specific mean of completed item responses
- Generate the overall total score by summing all 28 item responses (including imputed data)

#### 49. PHQ-9 Scoring System

The PHQ-9 has nine questions; which are assigned values of 0, 1, 2, and 3 to the response categories of 'not at all', 'several days', 'more than half the days', and 'nearly every day', respectively. A total score ranging from 0 to 27 is then computed by summing together the assigned scores for the nine questions (Kroenke et al., 2001). The total score can also be categorised with respect to increasing depression: none (0 to 4), mild depression (5 to 9), moderate depression (10 to 14), moderately severe depression (15 to 19), and severe depression (20 to 27). A higher score indicates increasing severity of depression. For a PHQ-9 score to be classed as valid at least 7 items should be complete. The following recommended approach is adopted when two or fewer item are missing (Datadictionary nhs PHQ-9, 2016):

- Calculate the participant specific mean based on items with responses
- Impute the missing items with the patient specific mean of completed item responses
- Generate the overall total score by summing all 9 item responses (including imputed data)
- Round total score to the nearest integer

#### 50. GAD-7 Scoring System

For the seven questions; values of 0, 1, 2, and 3 are assigned to the response categories of 'not at all', 'several days', 'more than half the days', and 'nearly every day', respectively. A total score is then computed by summing together the assigned scores for the seven questions (Spitzer et al., 2006). A higher score indicates increasing severity of anxiety, which could also be classified as: no (0 to 4), mild (5 to 9), moderate (10 to 14), and severe anxiety (15 to 21) (Spitzer et al., 2006). The score is valid if at least 5 items are completed. If one or two items are missing, they are substituted by the mean score of the participant non-missing items (Datadictionary nhs GAD-7, 2016).

#### 51. SMAS Scoring System

This has a total of 30 questions across 6 subdomains (Talking initiative, Investment behaviour, Variety, Multifunctionality, Self-efficacy, Positive frame of mind). A total score is calculated by summing responses of all 30 items across the six subdomains (Schuurmans et al., 2005). Subdomain scores will not be classed as valid if more than one item is missing. Missing items will be imputed with the mean score of the completed items for that subdomain for the participant (Correspondence with N Steverink, publication pending). The composite total score will be calculated only if all subdomain scores are calculated. The total

score ranges from 30 to 175, where higher scores indicate greater self management abilities. Only the total score will be reported.

#### 52. GSE Scale Scoring System

All ten questions are assigned values of 1, 2, 3, and 4 to the corresponding categories 'not true at all', 'hardly true', 'moderately true', and 'exactly true', respectively. A total score with possible range of 10 to 40 is then calculated by summing all of the 10 questions. For the total score to be valid at least 7 items must be complete. The following recommended approach is adopted in the case of missing data when at least 7 items are complete (Fu-berlinde, 2016):

- Calculate the participant specific mean based on items with responses
- Impute the missing items with the patient specific mean of completed item responses
- Generate the overall total score by summing all 10 item responses (including imputed data)

A higher GSE total score indicates more self-efficacy (Schwarzer et al, 1995).

#### 53. Diener's Flourishing Scale Scoring System

A total score based on all 8 items, with varying responses from 1 to 7 is computed to give a possible range of scores of 8 to 56. The higher the score means that a participant has many psychological resources and strengths. The score will not be calculated if any item responses are missing.

#### 54. IADL Scoring System

The IADL consists of 8 items that receive a score of 0 or 1. The total score is calculated by summing the scores for each item, the total score ranges from 0-8. A lower score indicates a higher level of dependence. The IADL will not be calculated if any item responses are missing.

#### 55. MMSE Scoring System

The Mini Mental State Examination (MMSE) consists of 11 questions. The total score is calculated by summing the responses. The score ranges from 0-30. A score of 23 or lower is indicative of cognitive impairment. The MMSE will not be calculated if item responses are missing.

#### 56. EQ-5D-5L Scoring system

The EQ-5D-5L consists of 5 questions, the score is measured on a scale from -0.22 to 1.00 (good health). The score will not be calculated if any items are missing. The algorithm for scoring the EQ-5D-5L can be found in the Database Specification. The EQ-5D-5L may be collected over the telephone or face-to-face, the scoring and analysis will be consistent regardless of the method of collection but we will note the number of instances of telephone collection.

The EQ-5D VAS your health state today is measured on a scale from 0 (worst imaginable health state) to 100 (best imaginable health state).

#### 57. SCQ scoring system

The Sense of Competence Questionnaire score is calculated by summing the answers to the 27 questions. The total score ranges from 27 to 135. A higher score represents a person with a better sense of competence. The SCQ will not be calculated if any questions are missing.

#### 58. Implementation of the SAP

This SAP will be used as a work description for the statistician involved in the trial. All analyses will be performed by the same statistician (under the supervision of senior trial statistician Professor Stephen Walters) and consequently none of the investigators involved in the trial will perform any of the statistical analyses.

Initially, the data manager will provide blinded data for preliminary checks by the statistician. Due to the nature of the data (clustering in one arm only) it is not feasible for blinded randomisation codes to be released before database lock. Following database freeze, unblinded data will be delivered to the statistician to define analysis sets and test statistical programs. Any queries will be communicated to the study and data manager prior to database lock, and any changes to the database during this time will be documented. The database will be locked after agreement between the statistician, data manager and study manager. It is expected that no data amendments should be required following database lock. However if an amendment is required, the process is documented in CTRU SOP DM012.

59. Questionnaire summary table

Name	No. of items	Score range	Description	Interpretation of score
DEMQOL	29	28-112	Measures health related QoI in people with dementia	higher DEMQOL score means higher health related quality of life
PHQ-9	9	0-27	Measures severity of depression	higher score indicates increasing severity of depression
GAD-7	7	0-21	is a severity measure of generalised anxiety disorder	higher score indicates increasing severity of anxiety
IALD	8	0-8	measures independent living skills	higher score indicates lower level of dependence
GSE	10	10-40	measures general perceived self-efficacy	higher score indicates more self efficacy
Diener's Flourishing scale	8	56	measures psychological flourishing	higher score means participant has many psychological resources and strengths
SMAS	30	30-175	measures self management abilities	higher SMAS indicates greater self management abilities
SCQ	27	27-135	measures sense of competence in caregivers	higher score represents a person with a better sense of competence
EQ-5D-5L value index	5	(-0.224, 1)	Measure of health status. 5 domains include mobility, self-care, usual activities, pain/discomfort, and anxiety/depression.	A score of zero means death, 1 is full health, negative score is a state worse than death
EQ-5D-5L VAS	1	(0, 100)	Measure of health status.	A score of zero means worst health and 100 means best health.

## 60. Example Tables and Figures

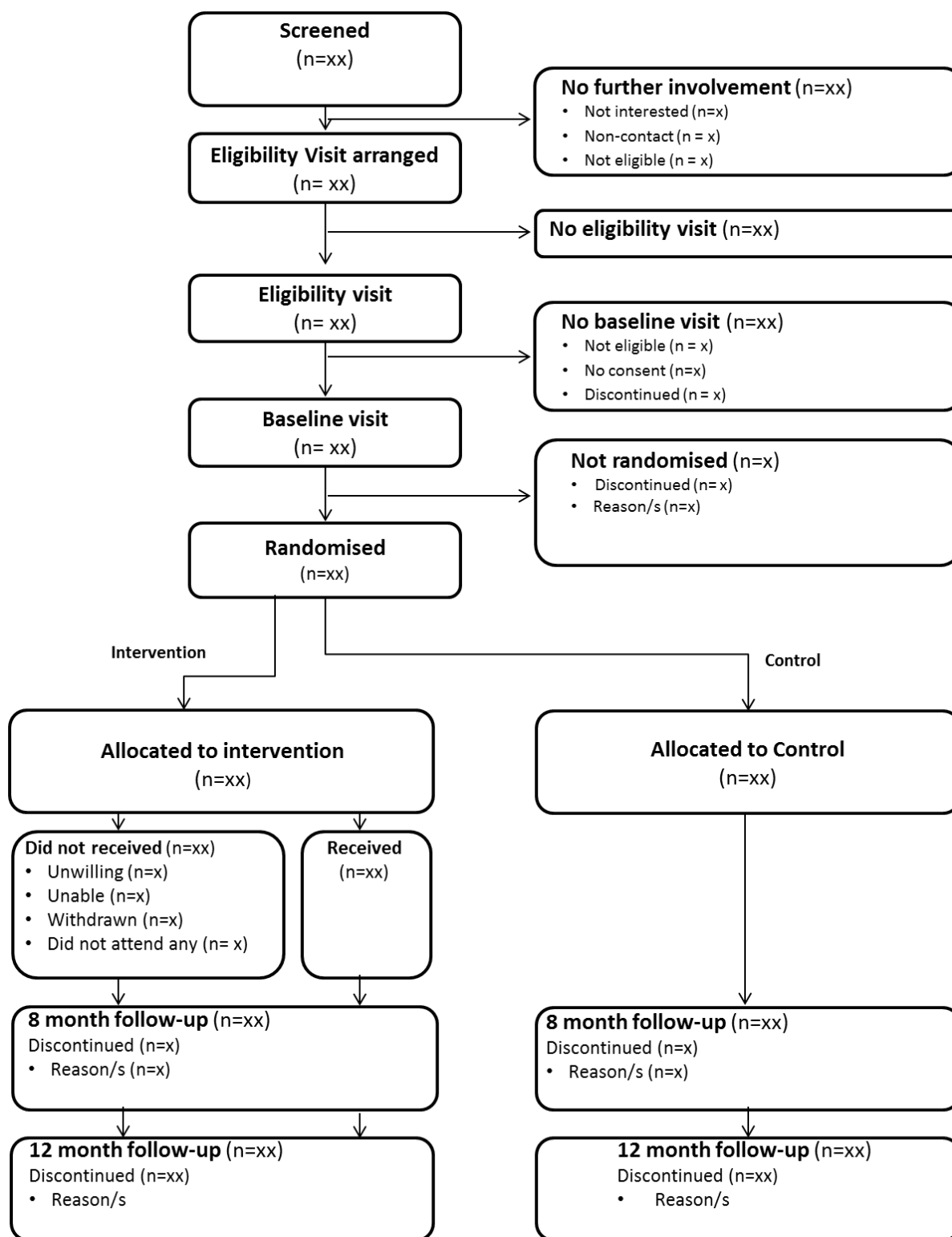


Figure 1 CONSORT flow diagram of participants through the trial

Table 4. Socio-demographic and characteristics of participants.

Variable	Scoring	Usual Care	JtD	All
		(n=xx)	(n=xx)	(N=xx)
Delivery site	Sheffield	xx(xx%)	xx(xx%)	xx(xx%)
	Bradford	xx(xx%)	xx(xx%)	xx(xx%)
	Leicester	xx(xx%)	xx(xx%)	xx(xx%)

	Hull	xx(xx%)	xx(xx%)	xx(xx%)
	York	xx(xx%)	xx(xx%)	xx(xx%)
	Halifax	xx(xx%)	xx(xx%)	xx(xx%)
	Nottingham	xx(xx%)	xx(xx%)	xx(xx%)
	Leeds	xx(xx%)	xx(xx%)	xx(xx%)
	Oldham	xx(xx%)	xx(xx%)	xx(xx%)
Sex	Male	xx(xx)	xx(xx)	xx(xx)
	Female	xx(xx)	xx(xx)	xx(xx)
Age (years)	Mean (SD)	xx(xx)	xx(xx)	xx(xx)
	Median (IQR)	xx(xx to xx)	xx(xx to xx)	xx(xx to xx)
	Min to Max	xx to xx	xx to xx	xx to xx
Ethnicity <sup>a</sup>	White <sup>b</sup>	xx(xx%)	xx(xx%)	xx(xx%)
	Mixed/multiple ethnic groups <sup>c</sup>	xx(xx%)	xx(xx%)	xx(xx%)
	Asian/Asian British <sup>d</sup>	xx(xx%)	xx(xx%)	xx(xx%)
	Black/African/Caribbean/Black British <sup>e</sup>	xx(xx%)	xx(xx%)	xx(xx%)
	Other ethnic group <sup>f</sup>	xx(xx%)	xx(xx%)	xx(xx%)
	Prefer not to say	xx(xx%)	xx(xx%)	xx(xx%)
Living with others	Yes	xx(xx%)	xx(xx%)	xx(xx%)
Live with (only yes above)	Spouse/partner	xx(xx%)	xx(xx%)	xx(xx%)
	Child/children	xx(xx%)	xx(xx%)	xx(xx%)
	Parent(s)	xx(xx%)	xx(xx%)	xx(xx%)
	Other	xx(xx%)	xx(xx%)	xx(xx%)

<sup>a</sup> Main ethnic groups could be collapsed depending on the observed distribution. <sup>b</sup> White: English/Welsh/Scottish/Northern Irish/British, Irish, Gypsy or Irish Traveller, and Any other White background; <sup>c</sup> Mixed/multiple ethnic groups: White and Black Caribbean, White and Black African, White and Asian, and Any other mixed/multiple ethnic groups background; <sup>d</sup> Asian/Asian British: Indian, Pakistani, Bangladeshi, Chinese, and Any other Asian background; <sup>e</sup> Black/African/Caribbean/Black British: African, Caribbean, and Any other Black/African/Caribbean/Black British background; <sup>f</sup> Other ethnic group: Arab, and Any other ethnic group.

Table 5: Baseline medical history of participants.

Variable	Scoring	Usual Care	JtD	All
		(n=xx)	(n=xx)	(N=xx)
Specific health problems	Stroke	xx(xx%)	xx(xx%)	xx(xx%)
	Diabetes	xx(xx%)	xx(xx%)	xx(xx%)
	Heart or chest problems	xx(xx%)	xx(xx%)	xx(xx%)
	Arthritis	xx(xx%)	xx(xx%)	xx(xx%)
	Sensory impairment	xx(xx%)	xx(xx%)	xx(xx%)
	Falls/dizziness/blackouts	xx(xx%)	xx(xx%)	xx(xx%)
	Any other	xx(xx%)	xx(xx%)	xx(xx%)
Type of dementia diagnosed	Alzheimer's	xx(xx%)	xx(xx%)	xx(xx%)
	Vascular dementia	xx(xx%)	xx(xx%)	xx(xx%)
	Mixed Alzheimer's/ Vascular dementia	xx(xx%)	xx(xx%)	xx(xx%)
	Other	xx(xx%)	xx(xx%)	xx(xx%)
	Not known	xx(xx%)	xx(xx%)	xx(xx%)

Table 6: Baseline characteristics and quality of life of participants for the ITT set.

Variable	Scoring	Usual Care	JtD	All
		(n=xx)	(n=xx)	(N=xx)
MMSE (total score) <sup>a</sup>		(n=xx)	(n=xx)	(N=xx)
	Mean(SD)	xx(xx)	xx(xx)	xx(xx)
	Media(IQR)	xx(xx to xx)	xx(xx to xx)	xx(xx to xx)
	Min to Max	xx to xx	xx to xx	xx to xx
MMSE cognitive impairment		(n=xx)	(n=xx)	(N=xx)
	Mild (18 to 23) <sup>**</sup>	xx(xx%)	xx(xx%)	xx(xx%)
	Normal (24 to 30)	xx(xx%)	xx(xx%)	xx(xx%)
DEMQOL (total score) <sup>b</sup>		(n=xx)	(n=xx)	(N=xx)
	Mean(SD)	xx(xx)	xx(xx)	xx(xx)
	Median(IQR)	xx(xx to xx)	xx(xx to xx)	xx(xx to xx)
	Min to max	xx to xx	xx to xx	xx to xx
DEMQOL Quality of life overall (Q29)		(n=xx)	(n=xx)	(N=xx)
	Very good	xx(xx%)	xx(xx%)	xx(xx%)
	Good	xx(xx%)	xx(xx%)	xx(xx%)
	Fair	xx(xx%)	xx(xx%)	xx(xx%)
	Poor	xx(xx%)	xx(xx%)	xx(xx%)
PHQ-9 (total score) <sup>c</sup>		(n=xx)	(n=xx)	(N=xx)
	Mean(SD)	xx(xx)	xx(xx)	xx(xx)
	Median(IQR)	xx(xx to xx)	xx(xx to xx)	xx(xx to xx)
	Min to Max	xx to xx	xx to xx	xx to xx
PHQ-9 Depression severity	None ( 0 to 4)	xx(xx%)	xx(xx%)	xx(xx%)
	Mild ( 5 to 9)	xx(xx%)	xx(xx%)	xx(xx%)
	Moderate (10 to14)	xx(xx%)	xx(xx%)	xx(xx%)
	Moderately severe (15 to 19)	xx(xx%)	xx(xx%)	xx(xx%)
	Severe (20 to 27)	xx(xx%)	xx(xx%)	xx(xx%)
GAD-7 (total score) <sup>d</sup>		(n=xx)	(n=xx)	(N=xx)
	Mean(SD)	xx(xx)	xx(xx)	xx(xx)
	Median(IQR)	xx(xx to xx)	xx(xx to xx)	xx(xx to xx)
	Min to Max	xx to xx	xx to xx	xx to xx
GAD-7 anxiety severity	Mild ( 0 to 5)	xx(xx%)	xx(xx%)	xx(xx%)
	Moderate (6 to10)	xx(xx%)	xx(xx%)	xx(xx%)
	Moderately severe (11 to 15)	xx(xx%)	xx(xx%)	xx(xx%)
	Severe (16 to 21)	xx(xx%)	xx(xx%)	xx(xx%)
IADL (total score) <sup>e</sup>		(n=xx)	(n=xx)	(N=xx)
	Mean(SD)	xx(xx)	xx(xx)	xx(xx)
	Median(IQR)	xx(xx to xx)	xx(xx to xx)	xx(xx to xx)
	Min to Max	xx to xx	xx to xx	xx to xx
GSE (total score) <sup>f</sup>		(n=xx)	(n=xx)	(N=xx)
	Mean(SD)	xx(xx)	xx(xx)	xx(xx)
	Median(IQR)	xx(xx to xx)	xx(xx to xx)	xx(xx to xx)
	Min to Max	xx to xx	xx to xx	xx to xx
Diener's Flourishing Scale (total score) <sup>g</sup>		(n=xx)	(n=xx)	(N=xx)
	Mean(SD)	xx(xx)	xx(xx)	xx(xx)
	Median(IQR)	xx(xx to xx)	xx(xx to xx)	xx(xx to xx)

SMA <sup>h</sup>	Min to Max	xx to xx	xx to xx	xx to xx
		(n=xx)	(n=xx)	(N=xx)
	Mean(SD)	xx(xx)	xx(xx)	xx(xx)
	Median(IQR)	xx(xx to xx)	xx(xx to xx)	xx(xx to xx)
EQ-5D-5L (value index) <sup>i</sup>	Min to Max	xx to xx	xx to xx	xx to xx
		(n=xx)	(n=xx)	(N=xx)
	Mean(SD)	xx(xx)	xx(xx)	xx(xx)
	Median(IQR)	xx(xx to xx)	xx(xx to xx)	xx(xx to xx)
EQ-5D VAS <sup>i</sup>	Min to Max	xx to xx	xx to xx	xx to xx
		(n=xx)	(n=xx)	(N=xx)
	Mean(SD)	xx(xx)	xx(xx)	xx(xx)
	Median(IQR)	xx(xx to xx)	xx(xx to xx)	xx(xx to xx)
	Min to Max	xx to xx	xx to xx	xx to xx
		(n=xx)	(n=xx)	(N=xx)
	Mean(SD)	xx(xx)	xx(xx)	xx(xx)
	Median(IQR)	xx(xx to xx)	xx(xx to xx)	xx(xx to xx)

<sup>a</sup> indicator of cognitive impairment <sup>b</sup> higher mean DEMQOL total scores means higher HRQoL; <sup>c</sup> a higher PHQ-9 total score indicates increasing severity of depression <sup>d</sup> a higher GDA-7 total score indicates increasing severity of anxiety <sup>e</sup> a lower IADL total score indicates a higher level of dependence; <sup>f</sup> higher GSE total score indicates more self-efficacy; <sup>g</sup> higher flourishing score represents a participant with many psychological resources and strengths; <sup>h</sup> a higher SMA composite total score relates to more positive well-being; <sup>i</sup> higher EQ-5D represents better health<sup>\*</sup>

Table 7: Baseline characteristics of participating supporters.

Variable	Scoring	Usual Care	JtD	All
		(n=xx)	(n=xx)	(N=xx)
Relationship with participants		(n=xx)	(n=xx)	(N=xx)
	Spouse	xx(xx%)	xx(xx%)	xx(xx%)
	Child	xx(xx%)	xx(xx%)	xx(xx%)
	Sibling	xx(xx%)	xx(xx%)	xx(xx%)
	Other family member	xx(xx%)	xx(xx%)	xx(xx%)
	Friend	xx(xx%)	xx(xx%)	xx(xx%)
	Neighbour	xx(xx%)	xx(xx%)	xx(xx%)
	Paid Carer	xx(xx%)	xx(xx%)	xx(xx%)
	Other	xx(xx%)	xx(xx%)	xx(xx%)
Live with participants	Yes	xx(xx%)	xx(xx%)	xx(xx%)
participants' first choice for support	Yes	xx(xx%)	xx(xx%)	xx(xx%)
Sex	Male	xx(xx)	xx(xx)	xx(xx)
	Female	xx(xx)	xx(xx)	xx(xx)
Age (years)	Mean (SD)	xx(xx)	xx(xx)	xx(xx)
	Median (IQR)	xx(xx to xx)	xx(xx to xx)	xx(xx to xx)
	Min to Max	xx to xx	xx to xx	xx to xx
SCQ (total score) <sup>a</sup>		(n=xx)	(n=xx)	(N=xx)
	Mean(SD)	xx(xx)	xx(xx)	xx(xx)
	Median(IQR)	xx(xx to xx)	xx(xx to xx)	xx(xx to xx)
	Min to max	xx to xx	xx to xx	xx to xx
PHQ-9 (total score) <sup>b</sup>		(n=xx)	(n=xx)	(N=xx)
	Mean(SD)	xx(xx)	xx(xx)	xx(xx)
	Median(IQR)	xx(xx to xx)	xx(xx to xx)	xx(xx to xx)
	Min to Max	xx to xx	xx to xx	xx to xx
EQ-5D-5L (value index) <sup>c</sup>		(n=xx)	(n=xx)	(N=xx)
	Mean(SD)	xx(xx)	xx(xx)	xx(xx)

	Median(IQR) Min to Max	xx(xx to xx) xx to xx	xx(xx to xx) xx to xx	xx(xx to xx) xx to xx
--	---------------------------	--------------------------	--------------------------	--------------------------

a), a higher SCQ score represents a better sense of competence; b) a higher PHQ-9 total score indicates increasing severity of depression; c) a higher EQ-5D-5L score represents better health.

Table 8. Exploratory effect of JtD intervention by subgroups: DEMQOL at 8 months.

Subgroup	Classification	Usual Care		JtD		Adjusted mean difference (95% CI) <sup>a</sup>	P-value <sup>b</sup>
		n	Mean(SD)	n	Mean(SD)		
Type of dementia	Alzheimer's	xx	xx(xx)	xx	xx(xx)	xx (xx to xx)	
	Vascular dementia	xx	xx(xx)	xx	xx(xx)	xx (xx to xx)	
	Mixed Alzheimer's/ Vascular dementia	xx	xx(xx)	xx	xx(xx)	xx (xx to xx)	
	Other	xx	xx(xx)	xx	xx(xx)	xx (xx to xx)	x.xxx
Presence of supporting carer	No	xx	xx(xx)	xx	xx(xx)	xx (xx to xx)	
	Yes	xx	xx(xx)	xx	xx(xx)	xx (xx to xx)	x.xxx

Usual care as the reference group; <sup>a</sup> adjusted for baseline DEMQOL (total score) and stratification site; <sup>b</sup> Overall interaction test. Higher mean DEMQOL total scores means higher HRQoL; SD=standard deviation; CI=Confidence Interval.

Table 9: Continuous baseline characteristics by group and missing data status

Variable	Summary Statistic	Completers			Non-completers		
		Usual Care (n=XX)	JtD (n=XX)	All (n=XX)	Usual Care (n=xx)	JtD (n=xx)	All (n=xx)
Age (yrs)	Mean(SD)	xx.x(xx.x)	xx.x(xx.x)	xx.x(xx.x)	xx.x(xx.x)	xx.x(xx.x)	xx.x(xx.x)
	Median(IQR)	xx.x(xx.x to xx.x)	xx.x(xx.x to xx.x)	xx.x(xx.x to xx.x)	xx.x(xx.x to xx.x)	xx.x(xx.x to xx.x)	xx.x(xx.x to xx.x)
MMSE	Mean(SD)	xx.x(xx.x)	xx.x(xx.x)	xx.x(xx.x)	xx.x(xx.x)	xx.x(xx.x)	xx.x(xx.x)
	Median(IQR)	xx.x(xx.x to xx.x)	xx.x(xx.x to xx.x)	xx.x(xx.x to xx.x)	xx.x(xx.x to xx.x)	xx.x(xx.x to xx.x)	xx.x(xx.x to xx.x)
DEMqoL	Mean(SD)	xx.x(xx.x)	xx.x(xx.x)	xx.x(xx.x)	xx.x(xx.x)	xx.x(xx.x)	xx.x(xx.x)
	Median(IQR)	xx.x(xx.x to xx.x)	xx.x(xx.x to xx.x)	xx.x(xx.x to xx.x)	xx.x(xx.x to xx.x)	xx.x(xx.x to xx.x)	xx.x(xx.x to xx.x)
PHQ-9	Mean(SD)	xx.x(xx.x)	xx.x(xx.x)	xx.x(xx.x)	xx.x(xx.x)	xx.x(xx.x)	xx.x(xx.x)
	Median(IQR)	xx.x(xx.x to xx.x)	xx.x(xx.x to xx.x)	xx.x(xx.x to xx.x)	xx.x(xx.x to xx.x)	xx.x(xx.x to xx.x)	xx.x(xx.x to xx.x)
GAD-7	Mean(SD)	xx.x(xx.x)	xx.x(xx.x)	xx.x(xx.x)	xx.x(xx.x)	xx.x(xx.x)	xx.x(xx.x)
	Median(IQR)	xx.x(xx.x to xx.x)	xx.x(xx.x to xx.x)	xx.x(xx.x to xx.x)	xx.x(xx.x to xx.x)	xx.x(xx.x to xx.x)	xx.x(xx.x to xx.x)
IADL	Mean(SD)	xx.x(xx.x)	xx.x(xx.x)	xx.x(xx.x)	xx.x(xx.x)	xx.x(xx.x)	xx.x(xx.x)
	Median(IQR)	xx.x(xx.x to xx.x)	xx.x(xx.x to xx.x)	xx.x(xx.x to xx.x)	xx.x(xx.x to xx.x)	xx.x(xx.x to xx.x)	xx.x(xx.x to xx.x)
GSE	Mean(SD)	xx.x(xx.x)	xx.x(xx.x)	xx.x(xx.x)	xx.x(xx.x)	xx.x(xx.x)	xx.x(xx.x)
	Median(IQR)	xx.x(xx.x to xx.x)	xx.x(xx.x to xx.x)	xx.x(xx.x to xx.x)	xx.x(xx.x to xx.x)	xx.x(xx.x to xx.x)	xx.x(xx.x to xx.x)
EQ-5D-5L (value index)	Mean(SD)	xx.x(xx.x)	xx.x(xx.x)	xx.x(xx.x)	xx.x(xx.x)	xx.x(xx.x)	xx.x(xx.x)
	Median(IQR)	xx.x(xx.x to xx.x)	xx.x(xx.x to xx.x)	xx.x(xx.x to xx.x)	xx.x(xx.x to xx.x)	xx.x(xx.x to xx.x)	xx.x(xx.x to xx.x)
Diener's Flourishing Scale	Mean(SD)	xx.x(xx.x)	xx.x(xx.x)	xx.x(xx.x)	xx.x(xx.x)	xx.x(xx.x)	xx.x(xx.x)
	Median (IQR)	xx.x(xx.x to xx.x)	xx.x(xx.x to xx.x)	xx.x(xx.x to xx.x)	xx.x(xx.x to xx.x)	xx.x(xx.x to xx.x)	xx.x(xx.x to xx.x)

Table 10: Categorical baseline characteristics by intervention group and missing data status.

Variable	Scoring	Completers			Non-completers		
		Usual Care (n=xx)	JtD (n=xx)	All (n=xx)	Usual Care (n=xx)	JtD (n=xx)	All (n=xx)
Sex	Male	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)
	Female	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)
Type of dementia diagnosed	Alzheimer's	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)
	Vascular dementia	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)
	Mixed Alzheimer's/ Vascular dementia	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)
	Other	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)
Live with others	Yes	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)	xx(xx.x%)
...	...	...	...	...	..	...	...

Table 11. Primary effectiveness analysis: DEMQOL total score at 8 months.

Primary outcome	Usual Care		JtD		Adjusted mean difference (95% CI) <sup>a</sup>	P-value <sup>a</sup>	Adjusted mean difference <sup>b</sup> (95% CI)	P-value <sup>b</sup>
	n	Mean(SD)	n	Mean(SD)				
DEMQOL total score	xx	xx(xx)	xx	xx(xx)	xx (xx to xx)	x.xxx	xx(xx to xx)	x.xxx

Usual care as the reference group; <sup>a</sup> adjusted for baseline DEMQOL (total score) and stratification site; <sup>b</sup> adjusted for baseline DEMQOL total score, stratification site, and additional covariates; age, sex, PHQ-9 (total score), and GAD-7 (total score). Higher mean DEMQOL total scores means higher HRQoL.

Table 12. Sensitivity analyses on the primary outcome: DEMQOL total score at 8 months.

Primary outcome: DEMQOL total score	Usual Care		JtD		Adjusted mean difference (95% CI) <sup>a</sup>	P-value <sup>a</sup>	Adjusted mean difference <sup>b</sup> (95% CI)	P-value <sup>b</sup>
	n	Mean(SD)	n	Mean(SD)				
MI	xx	xx(xx)	xx	xx(xx)	xx (xx to xx)	x.xxx	xx(xx to xx)	x.xxx
Regression imputation	xx	xx(xx)	xx	xx(xx)	xx (xx to xx)	x.xxx	xx(xx to xx)	x.xxx
CACE	xx	xx(xx)	xx	xx(xx)	xx (xx to xx)	x.xxx	xx(xx to xx)	x.xxx
MM	xx	xx(xx)	xx	xx(xx)	xx (xx to xx)	x.xxx	xx(xx to xx)	x.xxx

Usual care as the reference group; <sup>a</sup> adjusted for baseline DEMQOL (total score) and stratification site; <sup>b</sup> adjusted for baseline DEMQOL total score, stratification site, and additional covariates; age, sex, PHQ-9 (total score), and GAD-7 (total score). Higher mean DEMQOL total scores means higher HRQoL; SD=standard deviation, CI=Confidence Interval, CACE=Complier Average Causal Effect, MVI=Mean Value Imputation, MM=Mistimed Measurements.

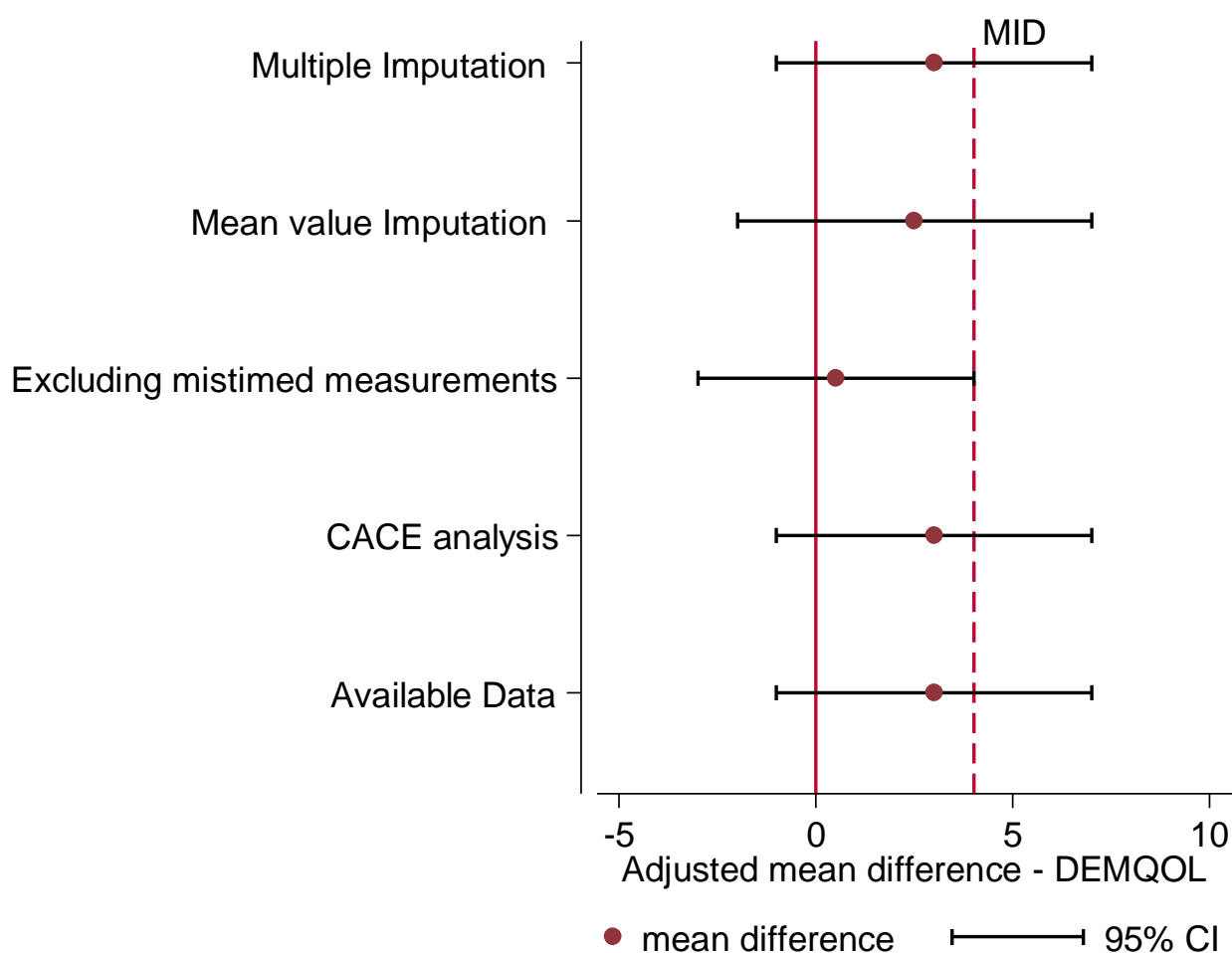


Figure 2 Forest plot of sensitivity analysis on primary outcome: DEMQOL at 8 months

Table 13. Exploratory effect of JtD intervention by subgroups: DEMQOL at 8 months.

Subgroup	Classification	Usual Care		JtD		Adjusted mean difference (95% CI) <sup>a</sup>	P-value <sup>b</sup>
		n	Mean(SD)	n	Mean(SD)		
Type of dementia	Alzheimer's	xx	xx(xx)	xx	xx(xx)	xx (xx to xx)	
	Vascular dementia	xx	xx(xx)	xx	xx(xx)	xx (xx to xx)	
	Mixed Alzheimer's/Vascular dementia	xx	xx(xx)	xx	xx(xx)	xx (xx to xx)	
	Other	xx	xx(xx)	xx	xx(xx)	xx (xx to xx)	x.xxx
Presence of supporting carer	No	xx	xx(xx)	xx	xx(xx)	xx (xx to xx)	
	Yes	xx	xx(xx)	xx	xx(xx)	xx (xx to xx)	x.xxx

Usual care as the reference group; <sup>a</sup> adjusted for baseline DEMQOL (total score) and stratification site; <sup>b</sup> Overall interaction test. Higher mean DEMQOL total scores means higher HRQoL; SD=standard deviation; CI=Confidence Interval.

Table 14. Effectiveness analysis: secondary outcomes at 8 and 12 months

Secondary outcome and timing	Usual Care		JtD		Adjusted mean difference (95% CI) <sup>a</sup>	P-value <sup>a</sup>
	n	Mean(SD)	n	Mean(SD)		
<b>At 8 Months</b>						
PHQ-9 (total score)	Xx	xx(xx)	xx	xx(xx)	xx (xx to xx)	x.xxx
GAD-7 (total score)	Xx	xx(xx)	xx	xx(xx)	xx (xx to xx)	x.xxx
IADL (total score)	Xx	xx(xx)	xx	xx(xx)	xx (xx to xx)	x.xxx
GSE (total score)	Xx	xx(xx)	xx	xx(xx)	xx (xx to xx)	x.xxx
Diener’s Flourishing Scale (total score)	Xx	xx(xx)	xx	xx(xx)	xx (xx to xx)	x.xxx
SMA						
Composite total score	Xx	xx(xx)	xx	xx(xx)	xx (xx to xx)	x.xxx
EQ-5D-5L (value index)	Xx	xx(xx)	xx	xx(xx)	xx (xx to xx)	x.xxx
EQ-5D-5L (VAS)	Xx	xx(xx)	xx	xx(xx)	xx (xx to xx)	x.xxx
<b>At 12 months</b>						
DEMQOL (total score)	Xx	xx(xx)	xx	xx(xx)	xx (xx to xx)	x.xxx
EQ-5D-5L (value index)	Xx	xx(xx)	xx	xx(xx)	xx (xx to xx)	x.xxx
EQ-5D-5L (VAS)	Xx	xx(xx)	xx	xx(xx)	xx (xx to xx)	x.xxx

Usual care as the reference group; <sup>a</sup> adjusted for baseline outcome measure under consideration and stratification site using a mixed effects linear regression model; SD=standard deviation; CI=Confidence Interval; Higher mean DEMQOL total scores means higher HRQoL; Higher mean flourishing score represents a participant with many psychological resources and strengths; a higher PHQ-9 total score indicates increasing severity of depression, a higher GAD-7 total score indicates increasing severity of anxiety; a lower IADL total score indicates a higher level of dependence; higher GSE total score indicates more self-efficacy; higher flourishing score represents a participant with many psychological resources and strengths; a higher SMA composite total score relates to more positive well-being; <sup>i</sup> higher EQ-5D-5L represents better health.

Table 15. Serious Adverse Events by treatment group

Serious Adverse Events	Intervention (n=XXX)	Control (n= XXX)	All (n= XXX)
Number (%) of participants who experienced ≥1 SAE	XXX (xx%)	XXX (xx %)	XXX (xx %)
Number of all SAEs (including repeated events)	XXX	XXX	XXX
Seriousness			
Death	xx (xx%)	xx (xx%)	xx (xx%)
Life threatening	xx (xx%)	xx (xx%)	xx (xx%)
Inpatient hospitalisation	xx (xx%)	xx (xx%)	xx (xx%)
Prolongs hospitalisation	xx (xx%)	xx (xx%)	xx (xx%)
Persistent or significant disability/incapacity	xx (xx%)	xx (xx%)	xx (xx%)
<i>Total</i>	xx (xx%)	xx (xx%)	xx (xx%)
Intensity			
Mild	xx (xx%)	xx (xx%)	xx (xx%)
Moderate	xx (xx%)	xx (xx%)	xx (xx%)
Severe	xx (xx%)	xx (xx%)	xx (xx%)
<i>Total</i>	xx (xx%)	xx (xx%)	xx (xx%)
Relationship to study intervention			
Unlikely	xx (xx%)		
Unrelated	xx (xx%)		
<i>Total</i>	xx (xx%)		

Table 16: Harms by treatment group: GAD-7 and PHQ-9

	Intervention (n=XXX)	Control (n= XXX)	All (n= XXX)
Number with symptoms of anxiety (GAD-7≥8)			
Baseline	XXX (xx%)	XXX (xx %)	XXX (xx %)
8 months	XXX (xx%)	XXX (xx %)	XXX (xx %)
12 months	XXX (xx%)	XXX (xx %)	XXX (xx %)
Number with symptoms of depression (PHQ-9≥10)			
Baseline	XXX (xx%)	XXX (xx %)	XXX (xx %)
8 months	XXX (xx%)	XXX (xx %)	XXX (xx %)
12 months	XXX (xx%)	XXX (xx %)	XXX (xx %)

## 61. References

- Abraham, I., Cherubini, A., Cozzolino, F., De Florio, R., Luchetta, M.L., Rimland, J.M., Folletti, I., et al. (2015), "Deviation from intention to treat analysis in randomised trials and treatment effect estimates: meta-epidemiological study.", *BMJ (Clinical Research Ed.)*, Vol. 350, p. h2445.
- Altman, D. G. (1990). *Practical statistics for medical research*. CRC press.
- Assmann, S.F., Pocock, S.J., Enos, L.E. and Kasten, L.E. (2000), "Subgroup analysis and other (mis)uses of baseline data in clinical trials.", *Lancet (London, England)*, Vol. 355 No. 9209, pp. 1064–9.
- Baldwin, S.A., Bauer, D.J., Stice, E. and Rohde, P. (2011), "Evaluating models for partially clustered designs.", *Psychological Methods*, Vol. 16 No. 2, pp. 149–65.
- Boutron I, Moher D, Altman DG, Schulz K, Ravaud P, for the CONSORT group. (2008) Methods and Processes of the CONSORT Group: Example of an Extension for Trials Assessing Nonpharmacologic Treatments. *Ann Intern Med*. W60-W67. PMID: 18283201
- Bsmsacuk. (2016). Bsmsacuk. Retrieved 15 November, 2016, from <https://www.bsms.ac.uk/research/cds/research/demqol.aspx>
- Brazier J, Ratcliffe J, Salomon J & Tsuchiya A (2007) *Measuring and Valuing Health Benefits for Economic Evaluation*. 1st Ed. Oxford: Oxford University Press.
- Datadictionary nhs PHQ-9. (2016). Datadictionarynhsuk. Retrieved 15 November, 2016, from [http://www.datadictionary.nhs.uk/data\\_dictionary/nhs\\_business\\_definitions/p/patient\\_health\\_questionnaire-9\\_de.asp?shownav=1](http://www.datadictionary.nhs.uk/data_dictionary/nhs_business_definitions/p/patient_health_questionnaire-9_de.asp?shownav=1)
- Datadictionary nhs GAD-7. (2016). Datadictionarynhsuk. Retrieved 15 November, 2016, from [http://www.datadictionary.nhs.uk/data\\_dictionary/nhs\\_business\\_definitions/g/generalised\\_anxiety\\_disorder\\_questionnaire\\_de.asp?shownav=1](http://www.datadictionary.nhs.uk/data_dictionary/nhs_business_definitions/g/generalised_anxiety_disorder_questionnaire_de.asp?shownav=1)
- de Boer, M.R., Waterlander, W.E., Kuijper, L.D.J., Steenhuis, I.H.M. and Twisk, J.W.R. (2015), "Testing for baseline differences in randomized controlled trials: an unhealthy research behavior that is hard to eradicate.", *The International Journal of Behavioral Nutrition and Physical Activity*, Vol. 12 No. 1, p. 4.
- Cuzick, J. (2005), "Forest plots and the interpretation of subgroups.", *Lancet (London, England)*, Vol. 365 No. 9467, p. 1308.
- Diener, E., Wirtz, D., Tov, W., Kim-Prieto, C., Choi, D. W., Oishi, S., & Biswas-Diener, R. (2010). New well-being measures: Short scales to assess flourishing and positive and negative feelings. *Social Indicators Research*, 97(2), 143-156.
- Flight, L., Allison, A., Dimairo, M., Lee, E., Mandefield, L., & Walters, S. J. (2016). Recommendations for the analysis of individually randomised controlled trials with clustering in one arm—a case of continuous outcomes. *BMC Medical Research Methodology*, 16(1), 165.
- Fu-berlinde. (2016). Fu-berlinde. Retrieved 15 November, 2016, from [http://userpage.fu-berlin.de/~health/faq\\_gse.pdf](http://userpage.fu-berlin.de/~health/faq_gse.pdf)
- Gold, D. A. (2012). An examination of instrumental activities of daily living assessment in older adults and mild cognitive impairment. *Journal of clinical and experimental neuropsychology*, 34(1), 11-34.
- ICH. (1998), ICH E9: Guidance on Statistical Principles for Clinical Trials, available at: [http://www.ich.org/fileadmin/Public\\_Web\\_Site/ICH\\_Products/Guidelines/Efficacy/E9/Step4/E9\\_Guideline.pdf](http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E9/Step4/E9_Guideline.pdf).
- Hancock, P., & Larner, A.J., (2009) "Clinical utility of Patient Health Questionnaire-9 (PHQ-9) in memory clinics" .*Int J Psychiatry ClinPract*; Vol. 13 No. 3, p.188-91.
- Kahan, B.C., Forbes, G., Ali, Y., Jairath, V., Bremner, S., Harhay, M.O., Hooper, R., et al. (2016), "Increased risk of type I errors in cluster randomised trials with small or medium numbers of clusters: a review, reanalysis, and simulation study", *Trials, BioMed Central*, Vol. 17 No. 1, p. 438.
- Knol, M., Groenwold, R. and Grobbee, D. (2012), "P-values in baseline tables of randomised controlled trials are inappropriate but still common in high impact journals", *European Journal of Preventive Cardiology*, Vol. 19 No. 2, pp. 231–232.
- Kroenke, K., Spitzer, R.L. and Williams, J.B. (2001), "The PHQ-9: validity of a brief depression severity measure.", *Journal of General Internal Medicine, Springer*, Vol. 16 No. 9, pp. 606–13.
- Mountain, G., Windle, G., Hind, D., Walters, S., Keertharuth, A., Chatters, R., Sprange, K., Craig, C., Cook, S., Lee, E., ..., Roberts, J.; (2017) A preventative lifestyle intervention for older adults (lifestyle matters): a randomised controlled trial. *Age Ageing* 1-8. doi: 10.1093/ageing/afx021

- Mulhern, B., Rowen, D., Brazier, J., Smith, S., Romeo, R., Tait, R., ... & Lamping, D. (2013). Development of DEMQOL-U and DEMQOL-PROXY-U: generation of preference-based indices from DEMQOL and DEMQOL-PROXY for use in economic evaluation. *Health technology assessment (Winchester, England)*, 17(5), 1-+.
- Pocock, S.J., Assmann, S.E., Enos, L.E. and Kasten, L.E. (2002), "Subgroup analysis, covariate adjustment and baseline comparisons in clinical trial reporting: current practice and problems.", *Statistics in Medicine*, Vol. 21 No. 19, pp. 2917-30.
- Peng, Y., Little, R. J., & Raghunathan, T. E. (2004). An extended general location model for causal inferences from data subject to noncompliance and missing values. *Biometrics*, 60(3), 598-607.
- Roberts, C., Batistatou, E. and Roberts, S.A. (2015), "Design and analysis of trials with a partially nested design and a binary outcome measure", *Statistics in Medicine*, No. April 2014, p. n/a-n/a.
- Satterthwaite, F.E. (1946), "An approximate distribution of estimates of variance components.", *Biometrics*, Vol. 2 No. 6, pp. 110-4.
- Schuurmans, H., Steverink, N., Frieswijk, N., Buunk, B.P., Slaets, J.P.J. and Lindenberg, S. (2005), "How to measure self-management abilities in older people by self-report. The development of the SMAS-30", *Quality of Life Research*, Kluwer Academic Publishers, Vol. 14 No. 10, pp. 2215-2228.
- Schwarzer, R., & Jerusalem, A., (1995) Generalized Self-Efficacy scale. In J.Weinman, S. Wright & M.Johnston. *Measures in health psychology: A user's portfolio. Causal and control beliefs*; 35-57. Windsor, UK: NFER-NELSON.
- Senn, S. (1994), "Testing for baseline balance in clinical trials.", *Statistics in Medicine*, Vol. 13 No. 17, pp. 1715-26.
- Smith, S.C., Lamping, D.L., Banerjee, S., Harwood, R., Foley, B., Smith, P., Cook, J.C., et al. (2005), "Measurement of health-related quality of life for people with dementia: development of a new instrument (DEMQOL) and an evaluation of current methodology.", *Health Technology Assessment (Winchester, England)*, Vol. 9 No. 10, pp. 1-93, iii-iv.
- Spitzer, R.L., Kroenke, K., Williams, J.B.W. and Löwe, B. (2006), "A brief measure for assessing generalized anxiety disorder: the GAD-7.", *Archives of Internal Medicine*, Vol. 166 No. 10, pp. 1092-7.
- Walters, S.J. (2010), "Therapist effects in randomised controlled trials: what to do about them.", *Journal of Clinical Nursing*, Vol. 19 No. 7-8, pp. 1102-12.
- Wang, R., Lagakos, S.W., Ware, J.H., Hunter, D.J. and Drazen, J.M. (2007), "Statistics in medicine--reporting of subgroup analyses in clinical trials.", *The New England Journal of Medicine*, Massachusetts Medical Society, Vol. 357 No. 21, pp. 2189-94.
- Wild, B., Eckl, A., Herzog, W., Niehoff, D., Lechner, S., Maatouk, I., ... & Löwe, B. (2014). Assessing generalized anxiety disorder in elderly people using the GAD-7 and GAD-2 scales: results of a validation study. *The American Journal of Geriatric Psychiatry*, 22(10), 1029-1038.
- Vernooij-Dassen, M., Kurz, X., Scuvee-Moreau, J., & Dresse, A. (2003). [The measurement of sense of competence in caregivers of patients with dementia]. *Revue d'epidemiologie et de sante publique*, 51(2), 227-235.

### **Box 1: Description of Journeying through Dementia Intervention**

The Journeying through Dementia intervention, summarised using the Template for Intervention Description and Replication (TIDieR) checklist<sup>1</sup>

#### **Rationale**

Journeying through Dementia is comprised of both group and individual sessions led by trained facilitators in an accessible community venue or out in the community e.g. shops, cafes, leisure centres. People with dementia do not need a supporting carer to take part. The intervention is readily tailored to meet individual needs.

#### **Materials**

- Copies of the manualised intervention<sup>14</sup> provided to all those involved in intervention delivery. Topics within the manual include understanding dementia, living with dementia, relationships, keeping mentally well, dementia and daily living, building and developing skills, keeping connected, and planning for the future.
- A minimum of two facilitators (senior support workers, allied health professions or nurses with appropriate experience) required to deliver every 12 week intervention at each site. (approximately one day commitment per facilitator throughout).
- A supervisor with relevant senior clinical experience at each site to provide weekly supervision to all facilitators.
- A two-day intervention delivery training workshop delivered to all facilitators and supervisors (following a manualised format) prior to intervention delivery.
- A local accessible community venue for the group element of each 12 week intervention.
- Small funding to support group and one-to-one work sessions out in the community.
- Proformas for facilitators to document all intervention sessions and send to all participants, written primarily for the benefit of participants and in accord with best practice guidance.
- A proforma to document the timing and content of facilitator supervision.

#### **Procedures**

- Those randomised to the intervention were invited to take part at each site when sufficient numbers reached for viable intervention delivery (8 -12 participants).
- Recruited participants were asked if they wished to nominate a carer to take part and if so that person was recruited to the study.
- Trained facilitators responsible for preparation of activities, delivery of group and one-to-one sessions, record keeping and participation in regular supervision.
- Intervention participants invited to take part in 12 weekly facilitated groups, with at least three sessions being held outside the meeting venue to promote putting learning into practice and mastery with support from others. Examples of possible topics for the group include navigating public

transport, identifying and using community resources such as leisure centres, and dining with others in the community. Group discussion of each topic the group selects is followed by facilitated exploration through in-venue practical sessions and didactic information if appropriate. Participants are then encouraged to take what they have learnt into community settings and work on their own challenges with both peer and facilitator support.

- Four one-to-one sessions scheduled across the 12 weeks, conducted between each participant and a facilitator in the home or local community to pursue individual goals they wish to work on. Examples of goals participants might identify include continuing to pursue activities or interests, identifying new pastimes, or keeping physically fit. Participants are enabled to undertake activities to take forward their goals with facilitator support. Therefore, individual sessions complement and reinforce the learning acquired through the group.
- Weekly facilitator supervision delivered by site-based supervisors, supported by clinicians from the research team.
- After all sessions, copies of a dementia friendly summary sent to each participant and to the research team by facilitators.
- Recruited carers invited to the first, sixth and final group session and one-to-one sessions if the recruited person with dementia agreed.

#### **Changes to protocol**

- The need to involve more sites and additional facilitators to meet the recruitment target, necessitated different formats for delivery of facilitator training in the later stages of the trial but content always followed the manualised format.
- Reviews of intervention delivery based on received records were circulated to all facilitators and supervisors to promote intervention adherence when the need became apparent.

1. Hoffmann TC, Glasziou PP, Boutron I, Milne R, Perera R, Moher D, et al. Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. *BMJ*. 2014 Mar 7;348:g1687.