THE LANCET

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: Underwood M, Lamb SE, Eldridge S, et al. Exercise for depression in elderly residents of care homes: a cluster-randomised controlled trial. *Lancet* 2013; published online May 2. http://dx.doi.org/10.1016/S0140-6736(13)60649-2.

Appendix

Exercise for depression in older care home residents. A cluster randomised controlled trial.

Authors: Underwood M, Lamb SE, Eldridge S, Sheehan B, Slowther A, Spencer A, Thorogood M, Atherton N, Bremner SA, Devine A, Diaz-Ordaz K, Ellard DR, Potter R, Spanjers K, Taylor SJC.

Table of Contents

Table of Figures

Table of Tables

Introduction

In this supplementary appendix we present additional material to help interpret the findings presented in the main paper. In part one, we define care homes and provide some international context. In part two, we present the characteristics of participants recruited after randomisation (Table S1). In part three we present data on depression in the care home population. In part four, we present information on analyses referred to, but not fully explained in the main paper: drug use, safety data, and the health economic analysis.

Part one: Care homes: definition and international context

The term care home in the context of the OPERA trial refers to UK residential and nursing homes. In the UK, care homes are typically small by international standards, with a typical median size of 30-40 places.[1] Residential homes provide accommodation and personal care, while nursing homes also provide 24 hour nursing for those with multiple frailties or specific nursing needs. Internationally, care home size tends to be higher, with unit sizes of 100-150 typical, for example in the Netherlands [2] and in the U.S. [3] In all countries, concern about standards of care is found, [4, 5] while dementia prevalence is high in all care home settings. Demographic change means that care home use is sure to increase despite efforts to focus care in the community, [6] including from a very low base in developing countries.[7]

Part Two: Characteristics of participants recruited after randomisation

Table S1. End of study characteristics of participants recruited to study after randomisation included in end-of-study cross-sectional analyses.

		Intervention			Control		
		% or mean	Number or SD	N	% or mean	Number or SD	N
N=127/132 providing usable outcome data at end of study				76			51
Demographic data							
Female		78%	59	76	80%	41	51
White		97%	74	76	98%	50	51
Taking antidepressants		24%	18	76	29%	14	49
Age (at recruitment)		87·9	8·0	76	87·9	7·2	51
Age left full-time education		14.7	2·0	63	15·4	2.0	39
Assessment data				76			49
Depressed (GDS-15 ≥5 or equivalent) [a]		46%	27	59	54%	23	43
GDS-15 score	**(0-15, 0 =best)**	5·0	3·8	59	5·2	3·5	43
MMSE score	**(0-30, 30 =best)**	20·4	6·4	55	16·8	6·6	41
SPPB score	**(0-12, 12 =best)**	1·8	2·3	61	1·5	2·1	44
EQ-5D score	**(-0.594 to 1, 1 =best)**	0·53	0·35	53	0·55	0·40	33
Pain today				56			40
None		71%	40		82%	33	
Mild-Moderate		29%	16		15%	6	
Severe		0%	0		3%	1	
Fear of falling		44%	25	57	30%	12	40
Proxy data				76			51
Social engagement				75			48
High		20%	15		6%	3	
Medium		21%	16		29%	14	
Low		59%	45		65%	31	
Barthel Index	**(0-100, 100 =best)**	61.6	26·6	66	57·3	28·8	42
Proxy EQ-5D	**(-0.594 to 1, 1 =best)**	0·51	0·38	75	0·44	0·37	50
Comorbidities recorded in home records							
Cancer		4%	3	76	10%	5	50
Stroke		20%	15	76	26%	13	50
Dementia		11%	8	76	33%	17	51
Depression		12%	9	76	16%	8	50
Anxiety		16%	12	76	16%	8	50
Osteoporosis		15%	11	76	10%	5	50
Chronic lung disease		14%	11	76	6%	3	50
Urinary incontinence		40%	30	75	67%	34	51

[a] If fewer than 15 items on the Geriatric Depression Scale were completed then we considered 'depression' present; if five positive responses, 13 or more items are completed, four when 12 or 11 items are completed and three when 10 items are completed. Nine or fewer items completed was considered missing.

Part Three: Depression

Data reveal that of the 891 residents recruited to the cohort study 374 (42%) were classified as depressed at baseline assessment (GDS-15) (Main paper, Table 2). Data from the care homes reveals that in home records only 180/878 (20·5%) had a recognisable diagnosis of depression (Main paper, Table 2) and of the depressed cohort only 92/365 (25%) had a depression diagnosis recorded (Table S2).

These data suggest that depression is a considerable problem in the elderly care home population and further exploration of the diagnosis and appropriate treatments is still needed.

Table S2 Co-morbidities in depressed cohort participants

N=374	Group			
	Intervention n=174		Control n=200	
	N	(%)	N	(%)
Cancer	7/ 169	4	20/ 195	10
Stroke	43/ 169	25	50/ 197	25
Dementia	37/ 169	22	42/ 197	21
Depression	45/ 169	27	47/ 196	24
Anxiety	34/ 169	20	39/ 195	20
Osteoporosis	14/ 169	8	18/ 195	9
Chronic lung disease	22/ 168	13	19/ 197	10
Urinary incontinence	99/ 169	59	111/ 197	56
Three individuals in this cohort did not consent to have data collected from their records				

Part Four: Drug use, mental health team visits, safety data, economic analysis

Antidepressant use

We collected data on all medications used over a one week period from the home records prior to randomisation, then repeated at three, six, nine and 12 month after randomisation. We extracted the exact drug name (as written in care home records), preparation, and dose used plus the number of times any medication was actually administered over a one week period. We then used the prescription cost analysis database to attach a code to each unique preparation used. [8] Using this methodology we calculated and estimated total amount used of each individual drug listed in the British National Formulary. [9] We included drugs in British National Formulary sections 4.3.1 Tricyclic and related antidepressant drugs, 4.3.3 Selective serotonin re-uptake inhibitors and 4.3.4 other antidepressant drugs in this analysis. Number of defined daily doses (DDDs) of antidepressants used was derived from the total mg of each antidepressant medication given in seven days from medical records data. These data were converted into defined daily doses using standard criteria (http://www.whocc.no/atc_ddd_index/). We excluded data on participants prescribed with low-dose (≤25 mg /day) amitriptyline which is typically used in pain management. We then interpolated forwards and backwards, in multiples of seven, to the midpoint date between each data collection visit to the homes. For residents who died, the interpolation value was scaled by the proportion of the time interval survived if this was less than a half. Otherwise a half was used, as was the case for residents who moved away from participating homes. For residents not taking any antidepressants in a time interval, their number of defined daily doses was set to zero for that interval. We used a linear mixed model to compare proportions taking antidepressant medication, a Mann-Whitney U test to compare number of defined daily doses of medication used and a mixed effects logistic regression to test for any differences in changes in who received antidepressant between baseline and one year. We did not find any significant differences in medication use (Table S3).

Table S3. Antidepressant use by those in cohort analysis during study

	Intervention			Control		
	% or mean and median	number or SD, 90th centile	N	% or mean and median	number or SD, 90th centile	N
Drugs Cohort Baseline (N = 869)			392			477
Antidepressant use in previous week [a]						
Baseline	28%	110	392	31%	150	477
3 months	28%	100	359	31%	131	427
6 months	29%	97	335	31%	124	395
9 months	29%	88	306	29%	105	364
12 months	29%	83	284	31%	103	333
Defined daily doses of medication over all follow-up [b]						
	59	123, 242	392	76	171, 282	477
Changes in antidepressant prescribing in those present at randomisation and end of follow-up [c]			282			333
At baseline and at 12m	26%	73		25%	84	
Only at 12m	4%	10		6%	19	
Only at baseline	4%	10		3%	11	
Never on drug	67%	189		66%	219	

[a] P-value from linear mixed model is 0·7779, [b] P-value from Mann-Whitney U test is 0·1677, [c] P-value from mixed effects logistic model is 0·2263

Mental health team visits

We collected data on visits by health professionals, including visits from the community mental health team, three, six, nine, and 12 months after randomisation from care home records. There were 168 such visits to 70 participants (range 1-17). In the control group 34/493 (7%) and in the intervention group 36/398 (9%) had one or more visit from the mental health team. We did not collect data on the reasons for these visits or whether the mental health team contact started before or after randomisation. These data do not suggest that our control intervention led to an increase in non-pharmacological treatments for depression.

Safety data

Everyone living in the intervention homes was exposed to the exercise intervention, including those who were not study participants. Safety monitoring, therefore, included all residents. We monitored any directly attributable adverse events that occurred during the exercise groups or during study assessments. We defined a directly attributable adverse event as an event that required external medical attention as consequence of participation in in the study.

To obtain an overall picture for each home during the intervention period we used the routinely collected data that care homes are required to keep on deaths and fractures. Care homes are specifically required to record and report deaths and serious injury, including falls and fractures under Regulation 37, Part V11 of The Care Home Regulations (2001). [10] We extracted pooled anonymous data unlinked to identity for the preceding three months from each home for all residents (participants and non-participants) three, six, nine and 12 months after randomisation.

At the end of the study Hospital Episodes Statistics Secondary Uses Service data were collected from each Primary Care Trust in which the study had been run. [11] The data included data for accident and emergency department attendances, hospital admissions, and outpatient attendances, with a diagnostic code.

Fractures of interest were pre-specified as peripheral fractures, defined as fractures not involving the spine. All spinal fractures were excluded as these are not strongly related to falls injury. We identified all fracture codes in The International Statistical Classification of Diseases and Related Health Problems tenth revision (ICD-10) [12] and then searched diagnostic descriptions from the inpatient, outpatient and accident and emergency departments for fracture diagnoses.

One fracture event may lead to multiple health service encounters, for example an accident and emergency department attendance, followed by an admission and a subsequent outpatient appointment. We, therefore, cross referred between datasets to identify linked episodes that represented a single fracture event. The quality of coding in accident and emergency data was poor; typically these reported an attendance accompanied by a broad, non-specific diagnostic description of 'dislocation/fracture/joint injury/amputation'. Out-patient data typically reported clinic attended but not the diagnosis. We were not able to link these to care home records data on fractures as these had been provided to us as pooled anonymous data. Where a participant had multiple peripheral fractures for one episode of care, this was treated as a single fracture event. We then allocated events to different levels of certainty that there had been a fracture:

- Confirmed fracture; definite fracture code identified, typically an inpatient admission.
- Probable fracture; an A&E attendance followed by an orthopaedic outpatient appointment.
- Potential fracture; an A&E attendance for an injury in which the term fracture appeared (amongst other diagnostic terms) in the diagnostic description with no further data provided.

During the study we used death data collected from homes to monitor safety. For the final end-of study analysis we used the more robust routine health service data. At the end of the study we collected data on date for whom we had permission to access their medical record data, from the National Health Service's Medical Record Information Service. [13]

Two of our safety analyses (on fracture and death rates) were based on data relating to study participants and conducted in a similar fashion to our primary and secondary analyses, using mixed effects logistic models. The third analysis was of fractures amongst all care home residents (aggregated fracture level). These fractures could not be attributed to individuals so we used a Poisson model at the care home level, adjusting for clustering with a random effect, and home location, home size, home type, mean age of home participants at pre-randomisation baseline, proportion female, percentage on antidepressants at pre-randomisation baseline, proportion of residents with moderate or severe cognitive impairment at pre-randomisation baseline assessment.

There were no directly attributable adverse events. There were no differences in deaths and fractures between the intervention and control homes (Table S4)

Table S4. Peripheral fractures and deaths

Outcome	Denominators	Intervention	Control	Effect estimate	95% CI	ICC	p
Number of peripheral fractures							
Care home data (all residents=2133)	973, 1160	48	42	IRR=1·14 [a]	(0·80 to 1·63)	0·03	0·5581
NHS data (all participants)	501, 553	33	39	Risk difference = -0·005			0·7648 [b]
Definite (fracture code identified, typically inpatient)		27	28				
Probable (A&E and outpatient appointment)		1	7				
Possible (A&E 'fracture' in diagnostic description)		5	4				
All cause mortality							
Participants present after randomisation	501, 553	119	122	OR=1·07 [c]	(0·78 to 1·48)	0·02	0·6745
Participants present at randomisation	402, 499			OR=1·08 [c]	(0·79 to 1·50)	0·001	0·6173

[a]Mixed effects Poisson model at the care home level, adjusting for mean age of residents, proportion of females, proportion of residents on antidepressants, proportion of residents with moderate to severe cognitive impairment, size of home, site, type of home and clustering with a random effect, IRR is incidence rate ratio. [b]χ^2 test. [c]Odds Ratio, derived from mixed effects logistic model, adjusting for mean age of residents, proportion of females, proportion of residents on antidepressants, proportion of residents with moderate to severe cognitive impairment, size of home, site and type of home.

Health economic analysis

Data for the economic evaluation covered a 12 month time period from when the home was randomised. A priori we chose to use proxy EQ-5D for this analysis as a poor rate of self-completion was anticipated. To ensure that quality adjusted life year (QALY) calculations were as accurate as possible we used multiple imputation to guard against any bias that may result from missing proxy EQ-5D scores. Since the study was cluster randomised rather than individual randomised, the data were multilevel. To adjust for this characteristic of the data, a multiple imputation model that accounted for clustering (by means of a random cluster effect) by care home was used to generate the missing proxy EQ-5D scores at each of the five time points (baseline and three, six, nine, and 12 months). The imputation model used the auxiliary variables home size, baseline age of resident and proxy EQ-5D scores at all five time points. We calculated the total utility for each participant from point estimates at baseline, three, six, nine and 12 months, using the 'area under the curve'. We assumed the utility for each person followed a linear trend line between these point estimates, [14] including those who died over the course of the study.

We collected data needed to calculate costs at the individual-level from routine Primary Care Trust data provided after the study was finished and through the care home records at three, six, nine, and 12 months. For resource use coming from care home records and community data provided by Primary Care Trusts, we obtained unit costs for the UK from published sources. [8, 15] Unit costs for medications were obtained from the Prescription Cost Analysis database for 2010, [8] and unit costs for hospital visits were obtained from the Primary Care Trust data. As the outpatient data had a large proportion of missing costs, and these missing data were more frequent with some Primary Care Trusts than with others, we imputed these data using the same methodology as we did for the missing EQ-5D scores before calculating the total outpatient cost. We estimated the cost of intervention from study data.

Because non-study participants were also exposed to the intervention and a proportion of the cost could also be attributable to them we used a weighted cost per resident. The fixed cost of providing the exercise classes was divided by the overall percentage of residents assessed as eligible to participate in the exercise sessions. This method reflected the whole home aspect of the intervention while removing the cost burden from those residents who were unlikely to receive much benefit from the programme due to communication difficulties or serious illness.

Our primary analysis was a cost-utility analysis over 12 months examining the cost per QALY gained for all participants who were assessed for proxy EQ-5D prior to randomisation and had Primary Care Trust data extracts. The clustering by home and the correlation between individual costs and outcomes needed to be considered in the cost-effectiveness analysis, so we used bivariate normal mixed models with cluster random effects. [16] We used multilevel multiple imputation of EQ-5D and outpatient data to handle the missing data, creating five imputed data sets of total costs and EQ-5D scores from which we calculated total costs and total QALYs. [17] Estimates of incremental costs and incremental QALYs obtained from each imputed set were combined using Rubin's rules to obtain multiple imputation estimates and standard errors. [18]

We plotted the cost-effectiveness plane, but calculated a mean incremental cost-effectiveness ratio only if it would be informative. For example, with data that show an insignificant quality-adjusted life year gain, it was more informative to report disaggregated data on costs and effects than report a mean incremental cost-effectiveness ratio. Since a definite threshold has proven difficult to set empirically, we planned to assess the cost-utility of the intervention using willingness to pay thresholds ranging between £0 and £40,000. [19] We did two sensitivity analyses; excluding high cost individuals (top 5%) and including costs from the societal perspective.

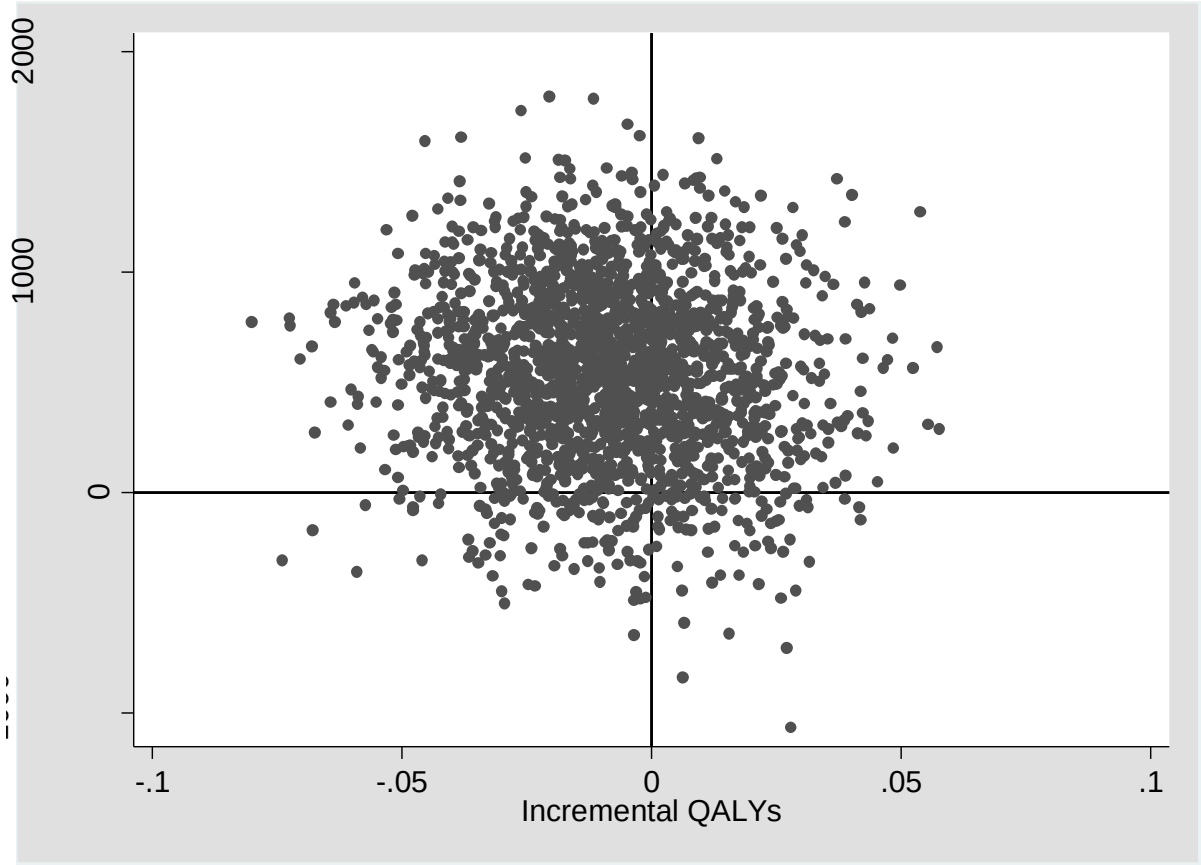
The base case analysis found the incremental QALY figure to favour the control arm of the study, the difference was negligible (0·0014) with a wide confidence interval (-0·073 to 0·070) (Table S5). The wide confidence interval means that we cannot formally conclude equivalence in QALYs since the limits of the 95% confidence interval include values where we might have concluded that the intervention was cost effective. In the cost-effectiveness plane the majority of results are in the north east quadrant (Figure S1). The sensitivity analyses gave similar results (Table S5)

The cost utility analysis shows that the intervention is dominated by the control demonstrating fairly conclusively that this is not a cost-effective intervention.

Table S5. Mean differences from the bivariate mixed models for the baseline and sensitivity analyses.

Measure		Mean difference (intervention – control)	95% CI
Base case analysis	Total costs £	374	-655 to 1404
	QALY	-0·001	-0·073 to 0·070
Excluding high cost participants	Total costs £	402	-224 to 1028
	QALY	-0·002	-0·076 to 0·072
Societal perspective	Total costs £	366	-664 to 1396
	QALY	-0·001	-0·073 to 0·070

Figure S 1. Cost-effectiveness plane generated from bootstrapped mean cost and QALY differences for residents over 12 months.



Positive values indicate that the intervention was more costly or more effective.

Table S6 CONSORT 2010 checklist of information to include when reporting a cluster randomised trial

Section/Topic	Item No	Standard Checklist item	Extension for cluster designs	Section of paper
Title and abstract				
	1a	Identification as a randomised trial in the title	Identification as a cluster randomised trial in the title	Title
	1b	Structured summary of trial design, methods, results, and conclusions (for specific guidance see CONSORT for abstracts)[i,ii]	See table 2	Abstract
Introduction				
Background and objectives	2a	Scientific background and explanation of rationale	Rationale for using a cluster design	Background
	2b	Specific objectives or hypotheses	Whether objectives pertain to the cluster level, the individual participant level or both	Methods
Methods				
Trial design	3a	Description of trial design (such as parallel, factorial) including allocation ratio	Definition of cluster and description of how the design features apply to the clusters	Methods
	3b	Important changes to methods after trial commencement (such as eligibility criteria), with reasons		N/A
Participants	4a	Eligibility criteria for participants	Eligibility criteria for clusters	Methods
	4b	Settings and locations where the data were collected		Methods & Results
Interventions	5	The interventions for each group with sufficient details to allow replication, including how and when they were actually administered	Whether interventions pertain to the cluster level, the individual participant level or both	Complex intervention only summarised
Outcomes	6a	Completely defined pre-specified primary and secondary outcome measures, including how and when they were assessed	Whether outcome measures pertain to the cluster level, the individual participant level or both	Methods
	6b	Any changes to trial outcomes after the trial commenced, with reasons		N/A
Sample size	7a	How sample size was determined	Method of calculation, number of clusters(s) (and whether equal or unequal cluster sizes are assumed), cluster size, a coefficient of intracluster correlation (ICC or *k*), and an indication of its uncertainty	Methods (sample size)
	7b	When applicable, explanation of any interim analyses and stopping guidelines		N/A
Randomisation:				
Sequence generation	8a	Method used to generate the random allocation sequence		Methods (recruitment and randomisation)
	8b	Type of randomisation; details of any restriction (such as blocking and block size)	Details of stratification or matching if used	as above
Allocation concealment mechanism	9	Mechanism used to implement the random allocation sequence (such as sequentially numbered containers), describing any steps	Specification that allocation was based on clusters rather than individuals and whether allocation concealment (if any) was at the cluster level, the individual	as above

		taken to conceal the sequence until interventions were assigned	participant level or both	
Implementation	10	Who generated the random allocation sequence, who enrolled participants, and who assigned participants to interventions	Replace by 10a, 10b and 10c	as above
	10a		Who generated the random allocation sequence, who enrolled clusters, and who assigned clusters to interventions	Methods
	10b		Mechanism by which individual participants were included in clusters for the purposes of the trial (such as complete enumeration, random sampling)	Methods (randomisation and recruitment)
	10c		From whom consent was sought (representatives of the cluster, or individual cluster members, or both), and whether consent was sought before or after randomisation	As above
Blinding	11a	If done, who was blinded after assignment to interventions (for example, participants, care providers, those assessing outcomes) and how		Methods (Masking)
	11b	If relevant, description of the similarity of interventions		Methods (Interventions)
Statistical methods	12a	Statistical methods used to compare groups for primary and secondary outcomes	How clustering was taken into account	Methods (Analysis)
	12b	Methods for additional analyses, such as subgroup analyses and adjusted analyses		N/A
Results				
Participant flow (a diagram is strongly recommended)	13a	For each group, the numbers of participants who were randomly assigned, received intended treatment, and were analysed for the primary outcome	For each group, the numbers of clusters that were randomly assigned, received intended treatment, and were analysed for the primary outcome	Results (Consort figures)
	13b	For each group, losses and exclusions after randomisation, together with reasons	For each group, losses and exclusions for both clusters and individual cluster members	Results
Recruitment	14a	Dates defining the periods of recruitment and follow-up		Results
	14b	Why the trial ended or was stopped		N/A
Baseline data	15	A table showing baseline demographic and clinical characteristics for each group	Baseline characteristics for the individual and cluster levels as applicable for each group	Results (tables)
Numbers analysed	16	For each group, number of participants (denominator) included in each analysis and whether the analysis was by original assigned groups	For each group, number of clusters included in each analysis	Results (Tables, consorts and text)
Outcomes and estimation	17a	For each primary and secondary outcome, results for each group, and the estimated effect size and its precision (such as 95% confidence interval)	Results at the individual or cluster level as applicable and a coefficient of intracluster correlation (ICC or k) for each primary outcome	Results (Tables, and text)
	17b	For binary outcomes,		N/A

		presentation of both absolute and relative effect sizes is recommended		
Ancillary analyses	18	Results of any other analyses performed, including subgroup analyses and adjusted analyses, distinguishing pre-specified from exploratory		N/A
Harms	19	All important harms or unintended effects in each group (for specific guidance see CONSORT for harms[iii])		N/A
Discussion				
Limitations	20	Trial limitations, addressing sources of potential bias, imprecision, and, if relevant, multiplicity of analyses		Discussion
Generalisability	21	Generalisability (external validity, applicability) of the trial findings	Generalisability to clusters and/or individual participants (as relevant)	Discussion and research in context
Interpretation	22	Interpretation consistent with results, balancing benefits and harms, and considering other relevant evidence		Discussion
Other information				
Registration	23	Registration number and name of trial registry		Title page
Protocol	24	Where the full trial protocol can be accessed, if available		References
Funding	25	Sources of funding and other support (such as supply of drugs), role of funders		Acknowledgements

References

1. Office Of Fair Trading. Care homes for older people in the UK: A market study. UK; 2005.
2. Schols JM, Crebolder HF, van Weel C. Nursing home and nursing home physician: the Dutch experience. J Am Med Dir Assoc. 2004; **5**(3): 207-12.
3. U.S. National Centre for Health Statistics. Health. USA; 2010.
4. Castle NG, Ferguson JC. What is nursing home quality and how is it measured? Gerontologist. 2010; **50**(4): 426-42.
5. Fahey T, Montgomery AA, Barnes J, Protheroe J. Quality of care for elderly residents in nursing homes and elderly people living at home: controlled observational study. BMJ. 2003; **326**(7389): 580.
6. Macdonald A, Cooper B. Long-term care and dementia services: an impending crisis. Age Ageing. 2007; **36**(1): 16-22.
7. Prince M. Care arrangements for people with dementia in developing countries. Int J Geriatr Psychiatry. 2004; **19**(2): 170-7.
8. PSU. Prescription Cost Analysis England 2010. 2011 [cited 2012 30.11.12]; Available from: http://www.ic.nhs.uk/services/prescribing-support-unit-psu
9. BNF. British National Formulary. 2011 12th January 2012 [cited; Available from: http://bnf.org/bnf/index.htm
10. HMSO. Social Care England Children and Young Persons England The Care Homes Regulations 2001. United Kingdom: HMSO; 2001. p. 17.
11. NHS. Using the Secondary Users Service. 2011 [cited 12th January 2012]; Available from: http://www.ic.nhs.uk/services/secondary-uses-service-sus/using-the-sus-service
12. WHO. ICD-10 Version:2010. 2010 5th January 2012 [cited 2011 12th January 2012]; Available from: http://apps.who.int/classifications/icd10/browse/2010/en
13. NHS. Medical Research Information Service. 2011 [cited 2011 12th January 2012]; Available from: http://www.ic.nhs.uk/services/medical-research-information-service
14. Richardson G, Manca A. Calculation of quality adjusted life years in the published literature: a review of methodology and transparency. Health Econ. 2004; **13**(12): 1203-10.
15. Curtis L. Unit costs of health & social care. 2010 [cited; Available from: http://www.pssru.ac.uk/archive/pdf/uc/uc2010/uc2010.pdf
16. Gomes M, Ng ES-W, Grieve R, Nixon R, Carpenter J, Thompson SG. Developing Appropriate Methods for Cost-Effectiveness Analysis of Cluster Randomized Trials. Medical Decision Making. 2012; **32**(2): 350-61.
17. Carpenter J, Goldstein H, Kenward M. REALCOM-IMPUTE Software for Multilevel Multiple Imputation with Mixed Response. Journal of Statistical Software. 2011; **45**(5): 1-14.
18. Rubin D. Multiple Imputation for Nonresponse in Surveys New York,: J. Wiley Sons; 1987.
19. Towse A. Should NICE's threshold range for cost per QALY be raised? Yes. BMJ. 2009; **338**: b181.