Cost-effectiveness of exercise as a therapy for Behavioural and Psychological Symptoms of dementia within the EVIDEM-E randomised controlled trial

Running head: The cost-effectiveness of exercise for BPSD

Keywords: Exercise, Dementia, Cost-effectiveness, NPI

Key points: The study suggests that exercise could potentially be a cost-effective intervention for outcomes measured by behavioural and psychological symptoms of dementia, but not when measured by QALYs.

Authors: Francesco D'Amico[1], Amritpal Rehill[1], Martin Knapp[1], David Lowery[2], Arlinda Cerga-Pashoja[2,3], Mark Griffin[3], Steve Iliffe[3] and James Warner[2,4]

Institutions: 1 Personal Social Services Research Unit, London School of Economics and Political Science; 2 Central & North West London NHS Foundation Trust (CNWL); 3 University College London (UCL) ; 4 Imperial College London

Contact details for corresponding author:

Dr F D'Amico, PSSRU LSE Health and Social Care, London school of Economics, Cowdray House, Houghton Street, London, WC2A 2AE, UK.

Tel: 020 7852 3714 email: Damico@lse.ac.uk

Sponsor: Central & North West London NHS Foundation trust

Sources of Support: This paper details independent research funded by a Programme Grant for Applied Research (RP-PG-0606-1005) granted by the National Institute for Health Research (NIHR). The views expressed within this article are those of the authors and not necessarily those of the NHS, NIHR or the Department of Health.

Abstract: 227 Word Count: 3413

Number of Tables: 5 Number of Figures: 1

ABSTRACT

Objective: Although available evidence is modest, exercise could be beneficial in reducing behavioural and psychological symptoms of dementia. We aim to evaluate the cost-effectiveness of a dyadic exercise regimen for individuals with dementia and their main carer as therapy for behavioural and psychological symptoms of dementia.

Methods: Cost-effectiveness analysis within a two-arm, pragmatic, randomised, controlled, single-blind, parallel-group trial of a dyadic exercise regimen (individually tailored, for 20-30 minutes at least five times per week). The study randomised 131 community-dwelling individuals with dementia and clinically significant behavioural and psychological symptoms with a carer willing and able to participate in the exercise regimen; 52 dyads provided sufficient cost data for analyses.

Results: Mean intervention cost was £284 per dyad. For the sub-sample of 52 dyads, the intervention group had significantly higher mean cost from a societal perspective (mean difference £2728.60, p=0.05), but costs were not significantly different from a health and social care perspective. The exercise intervention was more cost-effective than treatment as usual from both societal and health and social care perspectives for the measure of behavioural and psychological symptoms (Neuro Psychiatric Inventory). It does not appear cost-effective in terms of cost per quality-adjusted life year gain.

Conclusions

The exercise intervention has the potential to be seen as cost-effective when considering behavioural and psychological symptoms, but did not appear cost-effective when considering Quality Adjusted Life Year gains.

INTRODUCTION

Exercise has been suggested as a potential risk-reduction factor for dementia (Norton *et al.,* 2014; Ngandu *et al.,* 2015), and has also been suggested as a suitable intervention to address behavioural and psychological symptoms of dementia (BPSD) for people with established dementia (Brodaty and Arasaratnam, 2012). The EVIDEM-E trial investigated the effect of an individually tailored walking regimen for 131 dyads (pairs of people with dementia and their carer) on BPSD experienced by these individuals (Lowery *et al.,* 2014). The aim of this paper is to estimate the cost-effectiveness of this exercise intervention.

Non-pharmacological approaches are often effective, and may also be preferable alternatives to pharmacological interventions for BPSD. Exercise is potentially a particularly attractive option given its simplicity and potential positive spill-over effects when considering physical health. Nevertheless, it is important to consider the cost-effectiveness of any intervention, in order to ensure resources are being used efficiently, and that interventions are financially viable. This is particularly important for interventions aimed at assisting people with dementia given the context of an ageing population, which will place much greater future demands on health and social care systems that are already stretched (Prince *et al.,* 2014).

METHODS

Intervention design

The intervention delivered physical exercise in the form of 12-week individually tailored walking programme lasting for 20-30 minutes daily, designed to become progressively more intensive. Sessions were facilitated by a registered exercise professional qualified in instructing physical activity and exercise (National Vocational Qualification Level 3) and delivered to individuals within the intervention group in and around their own home. The exercise therapist progressively withdrew support over the first six weeks (and provided no support over weeks 7-12), with the expectation that the dyad would perform the exercise regimen regularly and independently at least five times per week. All participants were asked to record their daily activities throughout the 12 weeks of participation using a diary designed for the study. The intervention group diary contained an additional visual analogue scale, the Rating of Perceived Exertion (RPE) (Borg, 1982). Participants were encouraged to extend (*or reduce in some circumstance*) the level of intensity to between 12 and 14 on the RPE scale.

Participants

Participants were community-dwelling, and lived in inner city, urban and semi-rural areas in and around London. Recruitment was performed from the North Thames Dementias and Neuro-Degenerative Diseases Research Network's dementia research register (NTDEMREG)

(Iliffe *et al.,* 2011), through self-referral; through primary clinical services or through specialist mental health services (e.g., memory assessment and community mental health).

Participants were eligible for inclusion if they had a clinical diagnosis of dementia (defined by ICD-10 Diagnostic Criteria for Research (DCR-10) (1992)), and one or more significant BPSD symptom defined by a Neuropsychiatric Inventory (NPI; minimum score of frequency = 2, and severity = 2) (Cummings *et al.,* 1994), excluding hallucinations or delusions. To be eligible, participants also required a carer who was willing and able to participate with the exercise regimen, and to complete a falls risk assessment.

A full description of participants can be found in (Lowery *et al.,* 2014).

Randomisation

Study participants were assigned randomly to one of two trial arms. The intervention group received the exercise regimen in addition to treatment as usual (TAU). The control group received only treatment as usual. Individuals were allocated to these groups in a 1:1 ratio using a computer algorithm. Trial participants, carers, the participant's GP and the therapist were not blinded to treatment allocation; however, other individuals involved with the trail were blinded until analyses were completed. A full discussion of randomisation within the trial can be found in the published trial protocol (Cerga-Pashoja *et al.,* 2010).

Measures of outcomes

All outcomes were measured at baseline, at 6 weeks and 12 weeks (Lowery *et al.,* 2014). For the effectiveness analyses (Lowery *et al.,* 2014), the primary outcome measure was a between-group difference of proportions of people with a reduction of three or more points on the composite NPI score (Cummings *et al.,* 1994) at 12 weeks, which was chosen as a clinically significant change in BPSD symptoms. Secondary outcome measures were: mean difference in scores at 12 weeks on the NPI, General Health Questionnaire (GHQ) (Goldberg, 1972), DemQOL-Proxy (DEMQOL) (Smith *et al.,* 2007), and Zarit Caregiver Burden Inventory (ZBI) (Bedard *et al.,* 2001).

For the cost-effectiveness analyses, the primary outcome measure was mean difference in NPI score. Secondary outcome measures were mean difference in ZBI, GHQ and DEMQOL-proxy score, as well as QALYs calculated using these DEMQOL-proxy scores (Mulhern *et al.,* 2013).

Resource use and cost measures

Data on care and support service utilisation were collected using an adapted version of the Client Service Receipt Inventory (CSRI) (Beecham and Knapp, 2001), completed by the carer. The CSRI was completed at baseline and 12 weeks, and on each occasion asked about service receipt retrospectively over the previous 3 months. Data were collected on health and social care services (hospital services, day services, and community health and social care services), equipment and adaptations, medication and unpaid carer inputs.

Whenever possible, unit costs were taken from the Personal Social Services Research Unit (PSSRU) compendium for 2011 (Curtis, 2011), and reflect long-run marginal opportunity costs. The British National Formulary database was consulted with regards to costs for medication. Where costs for equipment and adaptations to home were not available in the PSSRU compendium they were estimated from market sources. Although most unit costs were found at 2011 prices, where this was not possible, available figures were adjusted to 2011 prices. Where services or equipment would continue to provide a benefit for more than one year (e.g. adaptations to home), costs were annuitised using the HM Treasury recommended annual discount rate of 3.5%. Unpaid care costs were estimated using an hourly rate equal to the National Minimum Wage, under the assumption that this was the potential opportunity cost for the unpaid carers.

Cost-effectiveness analyses

The cost-effectiveness analyses were conducted from two perspectives: the health and social care (HSC) perspective and societal perspective. The main difference between these perspectives is that the latter includes costs for unpaid carer time. The primary cost-effectiveness analyses from each perspective compared the exercise regimen and control groups on mean cost and mean difference in composite NPI score. Secondary cost-effectiveness analyses compared the groups on cost and each of the following outcomes in turn: the ZARIT caregiver burden inventory (ZBI), DEMQOL-Proxy, General Health Questionnaire (GHQ) and a measure of quality-adjusted life years (QALYs) generated from DEMQOL-Proxy scores. Scores on the outcome variables for which lower scores show better

outcomes have been reversed in order to have a more intuitive interpretation for the economic analysis.

If the exercise regime is more effective (has superior outcomes) and less costly than usual care, then it is said to *strongly dominate* the control intervention. If the exercise regimen has worse outcomes and higher costs than control, then it is said to be *strongly dominated*. In other circumstances the decision about whether or not to choose one intervention over the other is not straightforward, and the decision-maker must judge the differences in outcomes and costs before choosing one intervention over the other. The value or weight attached to differences in outcomes will play a part in making this decision. In such cases, we would calculate the incremental cost-effectiveness ratio (ICER):

$ICER = \Delta C / \Delta E,$

ΔC is mean incremental cost – the difference between the exercise regimen and control – and ΔE is the corresponding mean incremental outcome. The ICER aids the decision making process by displaying the cost (or cost savings) per unit change in the outcome considered.

Each ICER was estimated using the SUR (Seemingly Unrelated Regression) model within STATA (Stata Corp., 2013). Each cost and outcome measure in turn was regressed on treatment allocation, controlling respectively for cost and that same outcome measure at baseline. Regression models were bootstrapped with 1,000 replications in order to address potential skewness within the data. Multiple imputation (using ten imputed datasets) was employed to deal with missing values in some outcomes and covariates.

The formula $NB = \lambda * \Delta E - \Delta C$ was used to calculate net benefits (NB), which in addition to using mean cost and outcome differences, used a range of hypothetical values of willingness-to-pay (λ) for an additional unit on a given outcome measure. Cost-effectiveness acceptability

curves (CEACs) were then plotted for the primary outcome (NPI) using the net-benefit values calculated for each value of willingness-to-pay within the range of £0 to £10,000. This showed the probability of the exercise regimen being cost-effective over other willingness-to-pay values considered.

Economic analysis was conducted using STATA 13.

Sensitivity analyses

As part of a sensitivity analysis, we controlled additionally for participant age at baseline, gender, ethnicity, marital status, education level, whether living in a care home, MMSE score at baseline, carer's age and gender when performing regression analyses on cost and outcome measures when estimating ICERs.

Results

Sample characteristics

One hundred and thirty-one participant dyads were randomized to the intervention and control groups, and 113 (89%) completed the trial. The two groups were similar with respect to mean age, type of dementia and other characteristics including outcome scores. Descriptive statistics for the baseline demographic and the outcomes are presented in (Lowery *et al.,* 2014).

Completed CSRIs were received from 74 dyads at baseline and 67 at 12-week follow-up. Depending on the outcome variable however, the matching sample for the economic analysis varied from between 49 and 52 dyads because of missing data on some measures. Multiple imputation techniques with chained equations, were used to estimate missing outcome and socio-demographic data (Rubin, 1987), yielding a subsample of 52 dyads (22 within the control group, 30 in the intervention group) which could be analysed. Hereafter, our findings on service use, costs and cost-effectiveness are based on the 52 dyads in this sub-sample, which is slightly less than half of the sample available for the main outcome analysis (116 dyads).

We tested whether the economic analysis subsample was different from the original sample of 131 dyads by using the Wilcoxon rank-sum test for continuous variables and the Fisher Exact Test (or the Pearson Chi-2 test when appropriate) for binary variables, as shown in table 1. We found no significant differences between the 'economic subsample' and the rest of the sample at confidence level of 95% (with the smallest p-value of 0.07 associated with the 'primary education or less' variable). Looking at the differences between the intervention and control groups within the 'economic subsample', again no significant differences were found, with the smallest p-value of 0.07 for MMSE score at baseline.

Table 1

Service receipt

Service utilisation rates are presented in table 2. At baseline, there were no significant differences in service use by treatment group. At 12-weeks follow-up, overall proportions remained fairly similar. The two groups diverged with respect to utilisation of hospital services, with the percentage utilising these in the control group increasing (to 73%) whereas the proportion using these in the intervention group decreased (to 47%). However, as was the case with every other service use category at follow-up and at baseline, the difference between trial arms was not significant at the 5% level (p=0.09).

Table 2

Costs

Average intervention cost per dyad

Total intervention cost was calculated by multiplying unit cost per visit (£60) or unit cost per phone call (£10) by number of contacts between the exercise professional and each dyad. Mean intervention cost per dyad was £284 (range £190 to £320).

Cost analyses

Table 3 displays baseline and 12-week follow-up service use costs. At baseline, the summary statistics suggest that a sizeable proportion of total health and social care (HSC) service costs are related to use of accommodation services. At baseline there was no significant difference between the groups in terms of mean accommodation service costs. Apart from

accommodation, hospital and community services displayed the highest aggregate costs. Total HSC costs were £3,205 and £2,655 for the intervention and control groups respectively; the difference was not significant. There was also no significant difference between the groups in terms of mean societal cost (control group, £9,218; intervention group, £11,017) inclusive of the provision of unpaid care (control, £6,563; intervention, £7,812).

At follow-up, there was a significant difference in the cost of medications between the control (mean £246.30) and the intervention group (£285.20) (p=0.04), but we are unclear why this occurred. There was no significant between group difference in the cost of unpaid care. Total societal costs, including intervention and provision of unpaid care, was £10,533 for the intervention group versus £7,805 for the control group. After adjustment for baseline covariates this difference was not significant.

Table 3

Outcomes

Table 3 also displays the differences in outcomes between the intervention and the control groups at both time points. The economic subsample showed a significant between-group difference in scores for the General Health Questionnaire (GHQ) at 12 weeks, with the intervention group showing better scores on average (18.0 vs 23.2) than the control group. It should be noted however that prior to multiple imputations, data on GHQ scores were

recorded on a slightly reduced economic subsample of 49 respondents, and this difference was not found to be significant within the entire sample (Lowery *et al.,* 2014). No other significant between group outcome differences were found, in particular at baseline, meaning that the effect of randomisation persisted when considering the economic subsample as opposed to the entire sample.

Cost-effectiveness

In Table 4 we report the incremental costs and incremental effects for the primary and secondary outcome measures.

From a HSC perspective, we found that incremental costs were negative, i.e. the intervention group had lower costs than the control group (by approximately £170) and incremental effects were positive, i.e. the intervention group achieved better outcomes. Although none of these differences was significant at the 5% level, the results suggest that the exercise regimen dominated treatment as usual. To examine that possibility further, given that there were wide confidence intervals on the incremental differences, we plotted the CEACs for the primary and secondary outcomes. The CEACs for the primary outcome (NPI) are shown in Figure 1. From a HSC perspective (the dashed line in Figure 1), the CEAC suggests that, at a willingness-to-pay of £500 per incremental improvement in outcome (i.e. per 1-point difference in NPI score), the exercise regimen is cost-effective with a probability higher than 80%.

From a societal perspective, the ICER was £421 per incremental difference in NPI score. If a reduction of at least three points in the NPI score can be considered clinically meaningful in this case (it is suggested that this may vary: http://npitest.net/faqs.html), then this result suggests that the cost of achieving a meaningful improvement is £1,263. Whether this would be seen as cost-effective is unclear, since there have not been discussions of cost-effectiveness thresholds for this outcome measure; as ever, it would be for decision-makers to make the judgement.

With respect to secondary outcomes from a societal perspective, the ICERs using DEMQOL-proxy and GHQ as outcome measures were £580 and £392, respectively. With ZBI as the outcome measure, the ICER was £1,055. Finally, the ICER for QALYs calculated using DEMQOL-proxy scores and societal weights was large, at £286,440.

Table 4

Figure 1

Sensitivity analyses

Table 5 presents results for a sensitivity analysis that, in addition to baseline costs and baseline outcomes, also controlled for participant age at baseline, gender, ethnicity, marital

status, education level, whether living in a care-home, MMSE score at baseline, carer age and gender. This analysis provides results which are consistent with the main analysis; in particular, for the primary outcome measure of change in NPI score, the ICER calculated was equivalent to that of the main analysis.

Table 5

Discussion

Summary of findings

Some of the authors of this paper previously reported on a randomized, controlled, single-blind, parallel-group trial in which no evidence was found to support the hypothesis that exercise tailored to participant-carer dyads would be clinically effective for the amelioration of behavioural and psychological symptoms in dementia (BPSD) (Lowery *et al.,* 2014).

In this paper we reported cost data and cost-effectiveness analyses for a subsample of people in that trial. Mean costs for the group following the exercise regimen were not significantly different over 12 weeks from costs for the control group receiving treatment as usual. This result was found whether a societal or health and social care (HSC) perspective was adopted.

With respect to differences in outcomes, a significant difference in GHQ score was discovered at 12 weeks in the economic subsample, but not in the full sample (Lowery *et al.,* 2014). The subsample also did not exhibit a significantly different ZBI score at 12 weeks, whereas this difference was significant (p=0.01) for the full sample.

Although differences in outcomes and costs were not statistically significant, the cost-effectiveness acceptability curves (from both HSC and societal perspectives) suggest a high probability of cost-effectiveness at values of willingness to pay for one-point improvements on the NPI as low as £1000. If a three-point NPI difference can be considered clinically meaningful (with suggestions that this varies: http://npitest.net/faqs.html), then decision-makers (such as commissioners) would need to be prepared to pay at least £3000 for an estimated 68% probability of cost-effectiveness (from a societal perspective) and an estimated 82% probability of cost-effectiveness (health and social care perspective), interpreted as the cost of achieving a clinically meaningful improvement. There is no established cost-effectiveness benchmark for NPI with which to compare these estimates in the way that there is for QALYs (e.g. stemming from the threshold recommended by National Institute for Health and Care Excellence (NICE) in England and Wales). But even with cost-effectiveness measured in terms of cost per QALY, whether an intervention is cost-effective still comes down to a *judgement* by the decision-maker as to the 'worth' of an outcome difference. By this same token, our estimated mean cost per QALY looks high relative to the £30,000 upper threshold generally associated with cost-effectiveness judgements by NICE. The sensitivity analysis (adjusting for baseline socio-demographic variables) generated findings consistent with the main analyses.

Comparison with similar studies

Many clinical studies have investigated whether physical exercise is linked to a reduction in dementia risk, with the meta-analysis by Hamer *et al.* (2009) concluding that physical exercise is a protective factor for older people against Alzheimer's disease and other dementias. Other systematic reviews of epidemiological studies draw similar conclusions (Sofi *et al.,* 2011; Taxeira, 2012). More recent meta-analyses demonstrate a positive impact of physical exercise on Alzheimer's disease risk-reduction (Beydoun *et al.,* 2014). The link between exercise and vascular dementia is less clear; meta-analysis by (Aarsland *et al.*, 2010) found high heterogeneity between studies and the presence of publication bias. While there have been a number of studies of the effectiveness of exercise in reducing decline in cognitive functioning, to our knowledge there has been no examination of its cost-effectiveness (Knapp *et al.,* 2013).

Strengths and limitations of the economic evaluation

The economic evaluation adopted a health and social care perspective for some analyses, and a societal perspective for others, ensuring that unpaid carer inputs were not overlooked, and to provide relevant results for a greater range of stakeholders. A range of outcome measures were examined, including QALYs generated from a dementia-specific measure. A limitation of the economic analyses was the sample size, which was 55% lower than the sample for the main outcomes analysis because service use data were not available for each participant. Even though the subsample for whom we had cost data was not

significantly different from the larger sample recruited into the trial, this loss of statistical power limits the conclusions that can be drawn. Some of the observed cost differences may have reached statistical significance with a larger sample. Another possible limitation is that only 85% of study participants provided information about support received from unpaid carers, yet we know that all sample members had a carer. It is notoriously difficult to estimate accurately the amount of time an unpaid carer spends with someone with dementia, which can have an impact on the cost estimated.

Implications for policy and practice

This trial demonstrated that regular simple exercise such as walking does not appear to be effective with respect to reducing the behavioural and psychological symptoms of dementia, it does appear to help reduce carer burden however (Lowery *et al.,* 2014). Our cost-effectiveness analyses demonstrate that the regular walking regimen would be likely to be seen as cost-effective when focusing on NPI as the outcome of interest, but not when looking at other outcome measures, for example QALYs.

References

Aarsland D, Sardahaee F S, Anderssen S, Ballard C, group tAsSSR. 2010. Is physical activity a potential preventive factor for vascular dementia? A systematic review. *Aging & mental health*; **14**(4); 386-95.

Bedard M, Molloy D W, Squire L, Dubois S, Lever J A, O'Donnell M. 2001. The Zarit Burden Interview: a new short version and screening version. *Gerontologist* **41**(5); 652-657.

Beecham J, Knapp M. 2001. Costing Psychiatric Interventions. In *Measuring Health Needs; 2nd ed.* Thornicroft G (eds). Gaskell, London; 200–24.

Beydoun M A, Beydoun H A, Gamaldo A A, Teel A, Zonderman A B, Wang Y. 2014. Epidemiologic studies of modifiable factors associated with cognition and dementia: Systematic review and meta-analysis. *BMC Public Health*; **14**:643.

Borg G. 1982. Psychophysical bases of perceived exertion. *Medicine and Science in Sports and Exercise* **14**(5); 377-381.

Brodaty H, Arasaratnam C. 2012. Meta-analysis of nonpharmacological interventions for neuropsychiatric symptoms of dementia. *Am J Psychiatry* **169**(9); 946-953.

Cerga-Pashoja A, Lowery D, Bhattacharya R, Griffin M, Iliffe S, Lee J, Leonard C, Ricketts S, Strother L, Waters F, Ritchie C W, Warner J. 2010. Evaluation of exercise on individuals with dementia and their carers: a randomised controlled trial. *Trials* **11**:53.

Curtis L. 2011. *Unit Costs of Health and Social Care 2011.* PSSRU, University of Kent.

Goldberg D P. 1972. *The detection of psychiatric illness by questionnaire.* London, Oxford University Press

Hamer M, Chida Y. 2009. Physical activity and risk of neurodegenerative disease: a systematic review of prospective evidence. *Psychol Med*; **39**(1); 3-11.

Iliffe S, Curry L, Kharicha K, Rait G, Wilcock J, Lowery D. 2011. Developing a Dementia Registry: a descriptive case study from North Thames DeNDRoN and the EVIDEM programme. *BMC Medical Research Methodology* **11**(1): 9.

Knapp M, Iemmi V, Romeo R. 2013. Dementia care costs and outcomes: a systematic review. *International Journal of Geriatric Psychiatry*. **28**(6); 555-561.

Lowery D, Cerga-Pashoja A, Iliffe S, Thuné-Boyle I, Griffin M, Lee J, Bailey A, Bhattacharya R, Warner J. 2014. The effect of exercise on behavioural and psychological symptoms of dementia: the EVIDEM-E randomised controlled clinical trial. *International Journal of Geriatric Psychiatry.* **29**(8); 819–827.

Mulhern B, Rowen D, Brazier J, Smith S C, Romeo R, Tait R, Watchurst C, Chua K-C, Loftus V, Young T, Lamping D, Knapp M, Howard R, Banerjee S. 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*. **17**(5); 1-160.

Ngandu T, Lehtisalo J, Solomon A, Levälahti E, Ahtiluoto S, Antikainen R, Bäckman L, Hänninen T, Jula A, Laatikainen T, Lindström J, Mangialasche F, Paajanen T, Pajala S, Peltonen M, Rauramaa R, Stigsdotter-Neely A, Strandberg T, Tuomilehto J, Soininen H Kivipelto M. 2015. A 2 year multidomain intervention of diet, exercise, cognitive training, and vascular risk monitoring versus control to prevent cognitive decline in at-risk elderly people (FINGER): a randomised controlled trial. *Lancet* [online] **385**. Available at: http://dx.doi.org/10.1016/S0140-6736(15)60461-5 [Accessed: 15 June 2015].

Prince M, Knapp M, Guerchet M, McCrone P, Prina M, Comas-Herrera A, Wittenberg R, Adelaja B, Hu B, King D, Rehill A, Salimkumar D. 2014. *Dementia UK: Update*. London: Alzheimer's Society.

Rubin, D.B. 1987. *Multiple Imputation for Nonresponse in Surveys*. J. New York: John Wiley & Sons.

Smith S C, Lamping D L, Banerjee S, Harwood R, Foley B, Smith P, Cook J C, Murray J, Prince M, Levin E, Mann A, Knapp M. 2007. Development of a new measure of health-related quality of life for people with dementia: DEMQOL. *Psychological Medicine* **37**(5): 737-746.

Sofi F, Valecchi D, Bacci D, Abbate R, Gensini GF, Casini A, Macchi C. 2011. Physical activity and risk of cognitive decline: a meta analysis of prospective studies. *Journal of internal medicine* **269**(1): 107-17.

Stata Corp. 2013. *Stata: Release 13. Statistical Software*. College Station, Texas: StataCorp LP.

Teixeira C V, Gobbi L T, Corazza DI, Stella F, Costa J L, Gobbi S. 2012. Non-pharmacological interventions on cognitive functions in older people with mild cognitive impairment (MCI). *Archives of gerontology and geriatrics* **54**(1): 175-180.

World Health Organisation. 1992. The ICD-10 Classification of Mental and Behavioural Disorders: Diagnostic criteria for research. *F00 - F09 Organic, Including Symptomatic, Mental disorders: Dementia*. Geneva.

Table 1: Economic analyses subsample descriptive statistics

Variable	Treatment group						Control vs. Intervention (economic subsample)	Economic subsample vs. rest of the sample
	Control		Intervention		Total			
	n	Mean ± S.D. /Cases (%)	n	Mean ± S.D. /Cases (%)	n	Mean ± S.D. /Cases (%)	Wilcoxon / Chi2 test p-value	Wilcoxon / Chi2 test p-value
Age	22	78.4 ± 9.07	30	78.6 ± 7.6	52	78.5 ± 8.2	0.66	0.91
Carer's age	19	62.2 ± 17.11	25	63.6 ± 16.4	44	63.0 ± 16.5	0.88	0.96
MMSE	22	17.5 ± 8.17	30	13.6 ± 7.4	52	15.3 ± 7.9	0.07	0.58
Gender: female	22	13 (59.1)	30	16 (53.3)	52	29 (55.8)	0.78	0.89
Carer's gender: female	22	15 (68.2)	30	20 (66.7)	52	35 (67.3)	1.00	0.90
Ethnicity: white	22	18 (81.8)	30	27 (90.0)	52	45 (86.5)	0.44	0.18
Married or in a civil partnership	22	15 (68.2)	30	22 (73.3)	52	37 (71.2)	0.76	0.62
Living alone	22	7 (31.8)	30	8 (26.7)	52	15 (28.8)	0.76	0.66
Paid carer	22	1 (4.5)	30	2 (6.7)	52	3 (5.8)	1.00	1.00
Living in a care-home	22	2 (9.1)	30	4 (13.3)	52	6 (11.5)	1.00	0.98
Primary education or less	21	1 (4.8)	28	1 (3.6)	52	2 (3.8)	1.00	0.07
Further education	21	2 (9.5)	28	5 (17.9)	52	7 (13.5)	0.68	0.18
Dementia severity: mild	22	11 (50.0)	28	8 (28.6)	50	19 (38.0)	0.15	0.67
Dementia severity: moderate	22	7 (31.8)	28	10 (35.7)	50	17 (34.0)	1.00	0.78
Dementia severity: marked	22	4 (18.2)	28	10 (35.7)	50	14 (28.0)	0.22	0.86

Note: Wilcoxon test was performed for continuous variables (age, MMSE, etc.), while Chi^2 test was performed for categorical variables (gender, ethnicity, marital status, etc.).

Table 2: Economic analyses subsample service utilisation patterns

Variable	Treatment group – number and percentage using services						Fisher Exact Test p-value
	Control		Intervention		Total		
Pre-baseline (3 months)							
Res. Care/ Accommodation	2	9%	4	13%	6	12%	1.00
Hospital services	14	64%	16	53%	30	58%	0.57
Community services	14	64%	23	77%	37	71%	0.36
Equipment and adaptations	12	55%	20	67%	32	62%	0.40
Day services	8	36%	13	43%	21	40%	0.78
Medications	22	100%	29	97%	51	98%	1.00
Unpaid care	19	86%	24	80%	43	83%	0.72
N	**22**	**100%**	**30**	**100%**	**52**	**100%**	
Follow-up (1-3 months)							
Res. Care / Accommodation	1	5%	3	10%	4	8%	0.63
Hospital services	16	73%	14	47%	30	58%	0.09
Community services	12	55%	21	70%	33	63%	0.38
Equipment and adaptations	12	55%	14	47%	26	50%	0.78
Day services	10	45%	12	40%	22	42%	0.78
Medications	22	100%	29	97%	51	98%	1.00
Unpaid care	18	82%	26	87%	44	85%	0.71
N	**22**	**100%**	**30**	**100%**	**52**	**100%**	

Table 3: Costs (over three months) incurred by the economic analyses subsample

Variable	Control (N=22)				Intervention (N=30)				Total (N=52)				Bootstrapped t-test
	Mean	Std. Dev.	Min	Max	Mean	Std. Dev.	Min	Max	Mean	Std. Dev.	Min	Max	Adjusted p-value
Pre-baseline costs (£)													
Accommodation	951.7	3,080.3	0	10,468	1,300.9	3,478.1	0	13,182	1,153.2	3,288.8	0	13,182	0.72
Hospital services	513.6	747.8	0	2,566.0	577.5	1,248.9	0	5,217.0	550.4	1,057.4	0	5,217.0	0.83
Community services	575.5	1,108.9	0	4,355.3	682.4	1,010.9	0	3,061.6	637.2	1,044.2	0	4,355.3	0.72
Equipment and adaptations	68.2	135.7	0	502.5	112.0	158.6	0	459.1	93.4	149.5	0	502.5	0.30
Day services	270.2	550.0	0	2,055.6	259.6	492.5	0	1,778.9	264.1	512.3	0	2,055.6	0.94
Medications	275.7	194.3	3	882.8	272.8	177.0	0	691.3	274.0	182.7	0	882.8	0.96
Total HSC	2,654.8	3,756.8	315.7	16,121	3,205.1	3,595.1	129.4	13,747	2,972.3	3,638.2	129.4	16,121	0.61
Unpaid care	6,563.3	4,953.9	0	15,366	7,812.1	6,273.3	0	24,870	7,283.8	5,733.3	0	24,870	0.42
Total Societal	9,218.2	5,647.8	465.9	21,113	11,017	5,719.5	691.3	25,145	10,256	5,704.5	465.9	25,145	0.24
Pre-baseline outcomes													
NPI	32.9	19.1	7.0	73.0	31.6	19.2	6.0	76.0	32.1	19.0	6.0	76.0	0.79
ZBI	17.0	7.7	2.0	32.0	19.0	9.0	3.0	36.0	18.1	8.5	2.0	36.0	0.37
DEMQOL	100.7	16.3	61.0	121.0	103.6	12.5	60.0	121.0	102.4	14.2	60.0	121.0	0.66
Utility score (based on DEMQOL)	0.71	0.15	0.44	0.94	0.72	0.15	0.51	0.94	0.72	0.14	0.44	0.94	0.16
GHQ	19.7	10.9	4.0	48.0	17.9	9.1	8.0	54.0	18.7	9.8	4.0	54.0	0.51
Follow-up 12 weeks costs (£)													
Accommodation	632.7	2,967.7	0	13,919	697.0	2,361.4	0	10,468	669.8	2,607.4	0	13,919	0.80
Hospital services	461.0	937.2	0	4,425.7	146.7	255.9	0	898.0	279.7	650.7	0	4,425.7	0.08
Community services	270.4	707.0	0	3,229.0	390.5	782.2	0	3,919.9	339.7	746.5	0	3,919.9	0.65
Equipment and adaptations	103.0	189.7	0	710.4	89.0	160.3	0	641.5	94.9	171.7	0	710.4	0.25
Day services	270.5	519.5	0	1,937.0	229.1	476.8	0	1,541.7	246.6	490.7	0	1,937.0	0.78
Medications	246.3	169.4	6	672.8	285.2	172.9	0	783.7	268.7	170.9	0	783.7	0.04
Total HSC	1,983.8	3,080.5	85.7	14,528	1,837.5	2,511.8	118.4	11,367	1,899.4	2,738.7	85.7	14,528	0.41
Intervention	0	0	0	0	284.0	43.2	190.0	320.0	163.8	145.4	0	320.0	-

Total HSC + intervention	1,983.8	3,080.5	85.7	14,528	2,121.5	2,509.7	417.5	11,627	2,063.2	2,737.5	85.7	14,528	0.76
Unpaid care	5,820.7	6,750.9	0	28,626	8,411.6	5,727.0	0	24,570	7,315.5	6,251.9	0	28,626	0.24
Total Societal + intervention	7,804.5	6,859.0	85.7	29,735	10,533	5,890.7	532.5	29,271	9,378.7	6,399.7	85.7	29,735	0.31
Follow-up 12 weeks outcomes													
NPI	27.6	16.7	4.0	62.0	22.5	18.7	0.0	75.0	24.7	17.8	0.0	75.0	0.32
ZBI	18.9	8.5	3.0	32.0	18.7	8.3	5.0	38.0	18.8	8.3	3.0	38.0	0.33
DEMQOL	101.3	13.5	67.0	118.0	105.6	9.7	82.0	121.0	103.8	11.5	67.0	121.0	0.25
Utility score (based on DEMQOL)	0.67	0.14	0.51	0.94	0.71	0.14	0.47	0.94	0.70	0.14	0.47	0.94	0.21
GHQ	23.2	10.1	7.0	42.0	18.0	7.7	5.0	37.0	20.2	9.1	5.0	42.0	0.05

Note: In relation to the outcome variables, for the control group, the sample size for ZBI was 21 at baseline and follow-up and 21 for GHQ only at follow-up, before the multiple imputation. For the intervention group, the original sample size for ZBI was 29 ad baseline and follow-up, and, only at follow-up, 29 for NPI and 28 for GHQ. In all the other cases, scores were recorded for the entirety of the economic subsample.

Table 4: Incremental costs and effects (controlling for baseline costs and outcome)

HSC perspective	**Incremental cost (£, 2010/2011) Mean; [95% bootstrap CI]**			**Incremental effect Mean; [95% bootstrap CI]**			**ICER**
0-12 weeks	**Mean**	**Upper CI**	**Lower CI**	**Mean**	**Upper CI**	**Lower CI**	-
NPI	-168.6	-1,232.8	895.6	4.07	-4.65	12.79	Intervention dominant
ZBI	-170.8	-1,234.6	893.1	1.54	-1.78	4.86	Intervention dominant
DEMQOL-Proxy	-165.6	-1,251.7	920.6	2.87	-1.94	7.68	Intervention dominant
QALY (DEMQOL-Proxy)	-169.7	-1,240.0	900.5	0.0055	-0.0031	0.0140	Intervention dominant
GHQ	-173.6	-1,235.8	888.6	4.19	-0.55	8.93	Intervention dominant
Societal perspective	**Incremental cost (£, 2010/2011) Mean; [95% bootstrap CI]**			**Incremental effect Mean; [95% bootstrap CI]**			**ICER**
0-12 weeks	**Mean**	**Upper CI**	**Lower CI**	**Mean**	**Upper CI**	**Lower CI**	
NPI	1,686.4	-1,407.1	4,780.0	4.01	-4.72	12.73	421
ZBI	1,641.1	-1,497.8	4,780.0	1.56	-1.75	4.86	1,055
DEMQOL-Proxy	1,635.9	-1,520.9	4,792.6	2.82	-1.97	7.61	580
QALY (DEMQOL-Proxy)	1,565.8	-1,592.6	4,724.2	0.0055	-0.0031	0.0140	286,440
GHQ	1,657.3	-1,471.8	4,786.4	4.23	-0.50	8.97	392

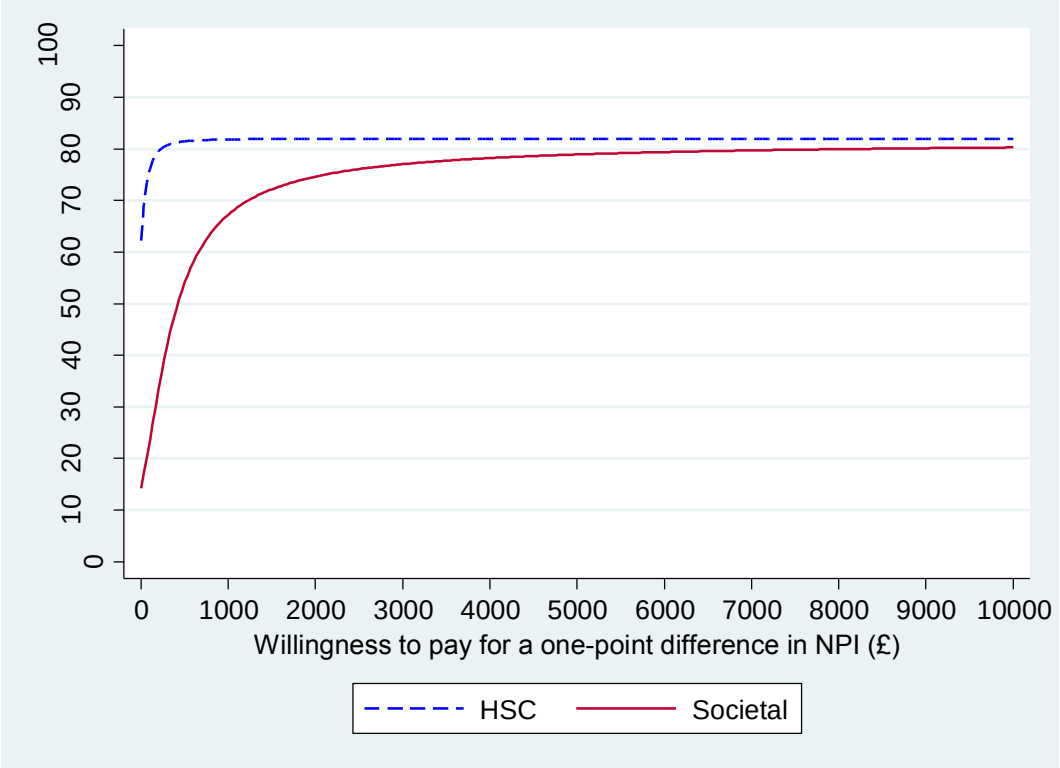
Note: Positive variations in the incremental effects represent improved outcomes. In order to obtain this, we reversed the scores for NPI, ZBI and GHQ. While dominance may be demonstrated, it must also be kept in mind that there were no significant differences in incremental costs and effects.

Table 5: Incremental costs and effects (controlling for socio-demographics, baseline costs and outcome)

HSC perspective	**Incremental cost (£, 2010/2011) Mean; [95% bootstrap CI]**			**Incremental effect Mean; [95% bootstrap CI]**			**ICER**
0-12 weeks	**Mean**	**Upper CI**	**Lower CI**	**Mean**	**Upper CI**	**Lower CI**	
NPI	-159.6	-1,267.8	948.5	2.46	-7.59	12.50	Intervention dominant
ZBI	-155.9	-1,254.7	942.9	0.56	-3.90	5.02	Intervention dominant
DEMQOL-Proxy	-156.5	-1,256.3	943.3	2.55	-3.32	8.41	Intervention dominant
QALY (DEMQOL-Proxy)	-156.7	-1,257.4	944.1	0.0066	-0.0026	0.0157	Intervention dominant
GHQ	-155.5	-1,250.6	939.5	4.00	-1.92	9.91	Intervention dominant
Societal perspective	**Incremental cost (£, 2010/2011) Mean; [95% bootstrap CI]**			**Incremental effect Mean; [95% bootstrap CI]**			**ICER**
0-12 weeks	**Mean**	**Upper CI**	**Lower CI**	**Mean**	**Upper CI**	**Lower CI**	
NPI	1,018.8	-2,331.0	4,368.7	2.42	-7.54	12.39	421
ZBI	992.5	-2,384.6	4,369.6	0.58	-3.86	5.02	1,711
DEMQOL-Proxy	978.1	-2,403.7	4,359.9	2.46	-3.40	8.33	397
QALY (DEMQOL-Proxy)	954.0	-2,444.5	4,352.4	0.0065	-0.0025	0.0155	146,437
GHQ	1,004.8	-2,349.2	4,358.8	4.07	-1.79	9.93	247

Note: Positive variations in the incremental effects represent improved outcomes. While dominance may be demonstrated, it must also be kept in mind that there were no significant differences in incremental costs or effects.

Figure 1: Cost-effectiveness acceptability curves: exercise regimen vs. usual care; health and social care and societal perspectives; effectiveness measured on the NPI scale



Declaration

Conflict of Interests: None to declare. The study sponsor role had no role in the study design, in the collection, analysis and interpretation of data, in the writing of the report and in the decision to submit this paper for publication.

Ethics: The study which collected the data on which this analysis was based obtained ethical approval to conduct research in collaboration with patients in the NHS and adhered to the stipulations of this approval. The ethical approval was granted by the Outer North East London Research Ethics Committee (ref 09/H0701/67). Permissions (CSP20305) were received for the study to recruit participants from the Central & North West London NHS Foundation Trust, Surry and Borders Partnership NHS Foundation Trust, West London Mental Health Trust, East London NHS Foundation Trust, and the following identification centres for participants: Camden & Islington NHS Foundation Trust, and the Barnet, Enfield and Haringey Mental Health NHS Trust.

Contributions of Authors: FD performed the cost-effectiveness analyses. FD and AR generated the economic cost variables necessary for the analyses, with FD, MK and AR overseeing the analyses and interpreting the data. DL, JW, SI, AP and MG each contributed significantly to the design and analyses of the original study, the results of which are used within these analyses. All aforementioned authors contributed to drafting this article and/or critically revising it where necessary. The final version of the article was approved by all authors, who each also take public responsibility for the content found within it.

Acknowledgements: The authors would like to thank each of the participants within the study- the individuals with dementia and their carers.