Evaluation of the Reminiscence Arts and Dementia: Impact on Quality of Life (RADIQL) programme in six care homes

Daniel Fujiwara
Ricky Lawton

October 2015

SIMETRICA

About the authors

Daniel Fujiwara (Bsc, Msc, Mphil) is Director of Simetrica. His research focuses on policy evaluation methods and techniques for valuing non-market goods. He has recently published guidelines on evaluation methods for the UK Government, including an update to the HM Treasury Green Book manual. Daniel previously led on cost-benefit analysis at the Department for Work and Pensions and has held senior economist positions at the Cabinet Office, the Ministry of Defence and the Ministry of Finance (Tanzania). In 2012 Daniel was awarded the *John Hoy Memorial Prize in Economics* for his contribution to policy evaluation in the UK Government.

In his current role at Simetrica Daniel has worked with and advised a wide range of OECD governments and international organisations, such as the United Nations, on policy evaluation.

Ricky Lawton (Bsc, Msc, PhD) is a Senior Economist at Simetrica. He specialises in quantitative and qualitative survey analysis, and measuring monetary values for environmental goods through preference methods, such as contingent valuation, and wellbeing valuation methods. Prior to joining Simetrica, Ricky worked in the Wellbeing Team at the Cabinet Office.

Acknowledgements

We would like to thank Age Exchange and in particular Belinda Sosinowicz and David Savill for their useful comments and inputs into the research. Data was collected by Dr. Frank Keating and Dr. Laura Cole at the Department of Social Work, Royal Holloway, University of London. Initial Dementia Care Mapping analysis was conducted by Robert Grant at the Centre for Health and Social Care Research, Kingston and St Georges University of London. All Intellectual property in the resulting outputs is the property of Age Exchange. The programme was funded with a grant of £600,000 by Guy's and St Thomas' Charity as part of their wider support to improve the health and wellbeing of the population of the London Boroughs of Lambeth and Southwark.

1. Introduction

It is predicted that 850,000 new people will be diagnosed with dementia in 2015, at an estimated cost of £26 billion to the public purse[1]. Research on the 400,000+ older people currently living in care homes finds evidence of rising depression and loneliness, and low levels of life satisfaction (Allen 2008; NICE 2014). Inactivity and low levels of engagement are cited as contributory factors, leading to loss of physical function, social isolation, behavioural symptoms and poor quality of life (NICE 2014).

Age Exchange responded to the needs presented by the growing numbers of people with dementia through a pioneering three-year research programme which tested the RADIQL intervention, a highly innovative person-centered form of creative care for people with complex health needs. RADIQL is a creative arts intervention that promotes communication and wellbeing by encouraging social inclusion, increased pleasure and cognitive stimulation (Zeilig et al. 2014). RADIQL aims to reduce levels of isolation, loneliness, unhappiness and depression for older people (Woods et al. 2005). It follows a personalised and relationship-centred reminiscence method centred on each individual's interests and life history set within a social context. The intervention used a variety of creative art forms and memory methods to connect people to their present, their social space, and to others around them.

Age Exchange tested the effectiveness of RADIQL for older people with dementia in six care homes in South London. The intervention has been independently evaluated by Royal Holloway University of London through quantitative and arts based research methods, using the widely recognised Dementia Care Mapping (DCM) tool. The aim of this study is to apply robust evaluation and cost effectiveness analysis that will enable Age Exchange to develop a high quality, evidence-based conceptual and pedagogic framework for social care workers, art practitioners and wider delivery in care homes for people with dementia.

The aim of this study was to provide a full assessment of the effectiveness of the RADIQL programme. We use **HM Treasury Green Book**[2] consistent methods in the evaluation.

The Green Book sets out the evaluation guidelines for all central government departments in the UK and it follows/aligns closely with best-practice guidelines issued by international organisations[3] and many other OECD governments[4]. The Green Book recommends **cost-benefit analysis** (**CBA**), where all of an intervention's costs and benefits are monetised and compared. And where this is not possible due to data constraints or difficulties in estimating monetary values the Green Book recommends

[1] www.alzheimers.org.uk/statistics
[2] https://www.gov.uk/government/publications/the-green-book-appraisal-and-evaluation-in-central-governent
[3] http://www.oecd.org/env/tools-evaluation/cost-benefitanalysisandtheenvironmentrecentdevelopments.htm
[4] http://www.fema.gov/government/grant/bca.shtm; http://ec.europa.eu/regional_policy/sources/docgener/guides/cost/guide2008_en.pdf

cost-effectiveness analysis (**CEA**), which is similar to CBA but it does not monetise the benefits of an intervention.

The two main outcomes that have been measured for the RADIQL programme are positive behaviour measures (measured through **Behaviour Category Codes** (**BCC**)) and a quality of life (QoL) measure (measured as a subjective **mood and engagement** (**ME**) score). These measures have not been monetised before in the health and care literature and the datasets do not provide information that would allow us to derive monetary values for these outcome measures through for example revealed preference methods and wellbeing valuation methods. Therefore, it was not possible to value the measured outcomes associated with the RADIQL programme and hence it is not possible to undertake a CBA study.

Instead, as stipulated in the Green Book, we use the data on the RADIQL programme to run a CEA study comparing the costs incurred to run the programme against BCC and ME outcomes. This provides a measure of relative programme effectiveness that can be compared against CEA results for other dementia related programmes.

This study is comprised of two main elements. We first estimate the impact of the RADIQL programme on the two outcome measures: BCC and ME. The focus of the study is on the impact of the programme over the period of data collection which was at baseline, at the 24 week end point and at a 40 week follow up. We estimate the impact of the programme through difference-in-difference analysis whereby the trends in outcomes for the programme group are compared to trends in outcomes for a control group.

The results of this impact analysis are then used in the CEA study, whereby the impact on outcomes is compared against the costs of administering the RADIQL programme.

2. Data

RADIQL observations involve the continual observation of participants with dementia over a sustained period of time carried out by a trained observer. Behaviour Category Codes (BCC) represent one of the 23 different domains of a participant's behaviour. The letters A-Z (excluding H, M, Z) represent the main behaviour observed within the time frame (Khuri 2010). In this analysis we recoded BCC codes on a 1-4 scale of lowest to high behaviour categories. Mood/engagement (ME) represents the quality of life of the participant associated with the BCC recorded in a time frame (on a -5 to 5 scale).

Around 100 care homes were identified through a desktop search of care homes in the London boroughs of Lambeth and Southwark. Thirty of those that conformed to the parameters of the research (i.e. provision of dementia services and residential care homes) were invited to participate. Of these 15 were suitable, having potential residents meeting the diagnosis of dementia living within a care home and not currently running a potentially conflicting programme. Final selection of 12 care homes was determined by availability of care homes during the study period, speed of recruitment and agreement to be part of the study. Seven care homes were from Lambeth and five from Southwark. Nine care homes were private and three were NHS run (GST and SLAM). Three of the twelve care homes were faith-based.

Six homes were assigned to receive the RADIQL intervention, and 6 were designated as control sites. Randomising the selection was not possible due to uncertainties around availability within the required timeframe, with a number of eligible care homes requesting to be placed in a hold group until they were certain of their ability to participate. In addition, the National Health Service (NHS) Ethics committee suggested that interventions be distributed among care homes on a like for like basis where possible.

Intervention sites received the RADIQL reminiscence treatment every week, with DCM observations once every three weeks. Data observations for BCC and ME were collected approximately every five minutes for a half hour before and a half hour after each treatment intervention (each RADIQL intervention lasts approximately one hour on average).

Baseline data was collected 3 to 4.5 weeks before the start of the intervention, varying slightly by each care home. The RADIQL intervention was performed every week, with the final intervention ('end point') performed in week 24. The end-point intervention ranged between 25-29 weeks from the baseline data collection, varying slightly by each care home (mean 27.3 weeks).

Individuals in the control group received no change to their normal care regime. Data on BCC and ME for the control group was captured at two points - at baseline and as close to 24 weeks as possible (mean 22.8 weeks, range 20-25 weeks). For the purposes of communicating results we report the control group end-point range as 20-30 weeks.

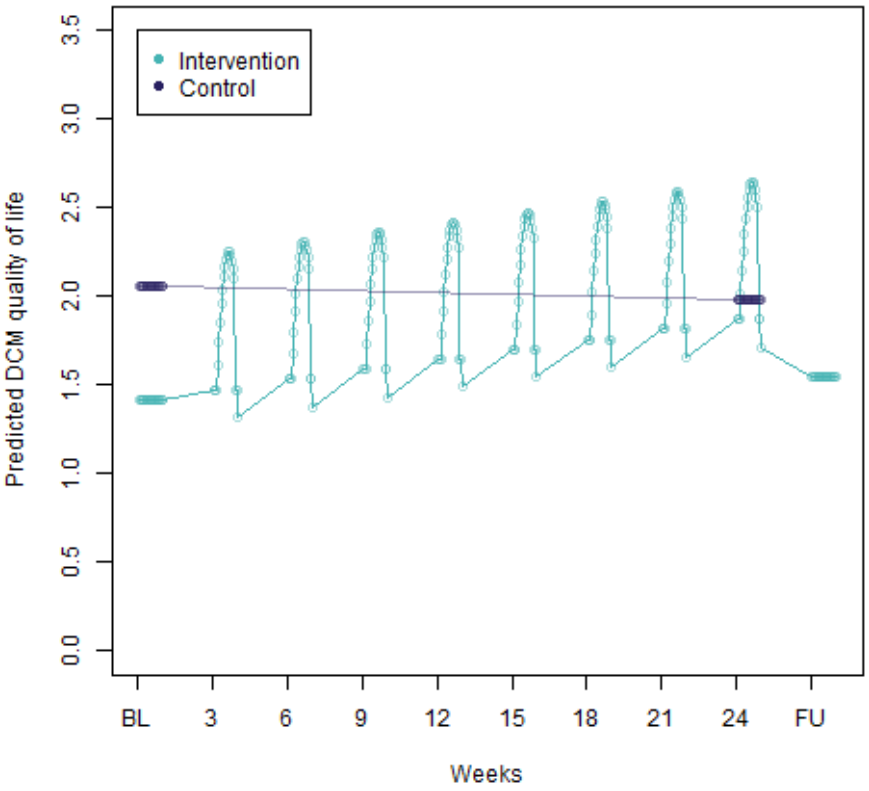
Follow-up observations were performed on the intervention group only at around 42 weeks after the baseline period (mean 42.2 weeks, ranging from 39-48 between

different homes). In order to compare with the control group, we extrapolate BCC and ME scores for the control group to the 42-week period by using the trend from baseline to the 24 week point. We note that this estimated trend line may overstate ME and BCC scores in the control group at week 42 because dementia conditions can deteriorate rapidly over longer time spans.

Sample size suffered from attrition over the course of the intervention. Total sample size at baseline period was n=68. The number of participants per home at baseline ranged from five to eight with the mean number being seven. Attrition over time reduced the final (end-point) sample to n=54. We exclude all observations where data is not available at end-point. Patient details on age, gender, ethnicity, date of admission, type of dementia, and medical observations were collected.

Figure 1 shows the data points that were collected for the treatment and control groups for the Dementia Care Mapping (DCM) quality of life measure (this is an aggregation of BCC and ME scores).

Figure 1 - Predicted Dementia Care Mapping (DCM) quality of life observations over time (provided in previous report)



We produced mean ME and BCC scores for each individual for the baseline (week 0), end point (week 24) and follow up (week 40+) periods (mean scores were averaged across all pre, during and post-intervention data observations). We recoded the ME variable on a 0-10 scale at two-unit intervals, where 0=extreme distress and 10=extreme positive mood and engagement. We assume that this scale is cardinal for the purposes of the regression analysis. We coded the BCC variable as set out in Table 1.

Table 1 - Positive behaviour code (BCC)

Level	Positive behaviour	Behaviour Category Code (BCC)
0	Lowest	N
1	Low	C, U, W
2	Medium	B
3	High	A, D, E, F, G, I, J, K, L, O, P, R, S, T, V, X, Y,

3. Methodology

We start with a simple test of the before-after change in outcomes for the treatment group. We tested for difference in outcome levels between baseline and end-point (20-30 weeks) and follow-up (40+ weeks) for the treatment group using paired t-tests.

We then performed a difference in difference (DiD) estimation model for the ME and BCC outcomes, which uses trends in the treatment and control groups controlling for background factors:

$$\Delta y_i = \alpha + \beta_1 T_i + \beta_2 X_i + + \varepsilon_i \qquad (1)$$

Where y is the outcome (ME or BCC) for individual i; T_i is the treatment variable, X_i is a vector of sociodemographic and health controls; ε_i is the error term representing unobserved factors; and: $\Delta y_i = y_{i1} - y_{i0}$

where $y_{i1}$is the outcome variable at post-intervention and y_{i0} is the outcome variable at pre-intervention.

Here β_1 is the DiD estimator for the impact of the RADIQL treatment on the outcome ME or BCC.

The DiD estimator of impact is unbiased if the trends for the control group (before and after the programme) provide a perfect representation of what would have happened for the treatment group had they not received the intervention (the counterfactual). This is the case even if pre-treatment outcomes are at different *levels* to being with. This cannot be tested but it can be assumed to hold if trends in the outcomes for a number of periods previous to the programme are the same for the treatment and control groups. This is known as the *parallel trends* assumption. Some studies require a *conditional parallel trends* assumption, which means that the parallel trends assumption holds after conditioning on some of the initial differences between the treatment and control groups.

Unfortunately in the present study the parallel trends assumption could not be tested because only one period of data was collected before the RADIQL programme started. Figures 2 and 3 respectively show the trends in ME and BCC scores from baseline (0 weeks) to end point (20-30 weeks). People who participated in (were selected into) the RADIQL programme had statistically significantly lower levels of ME and BCC than the control group participants. This, however, does not invalidate the DiD analysis since it is the trends rather than the levels of the outcomes that are important in DiD (the initial difference in the levels of the outcomes is controlled for in the analysis).

Figure 2 – Mean ME scores for intervention and control groups at intervention baseline (0 weeks) and end-point (20-30 weeks)

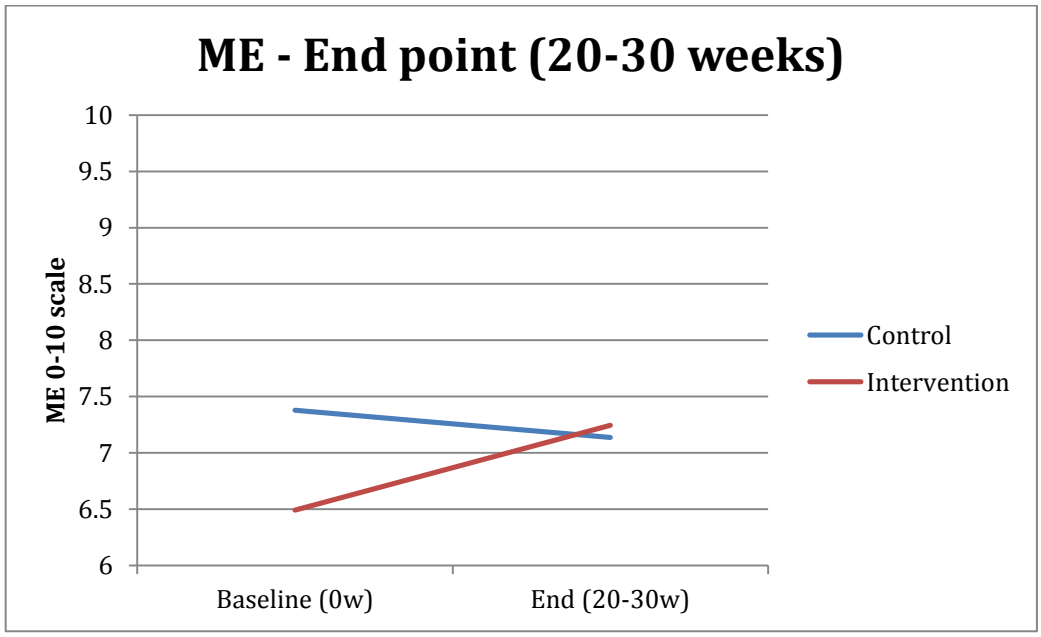
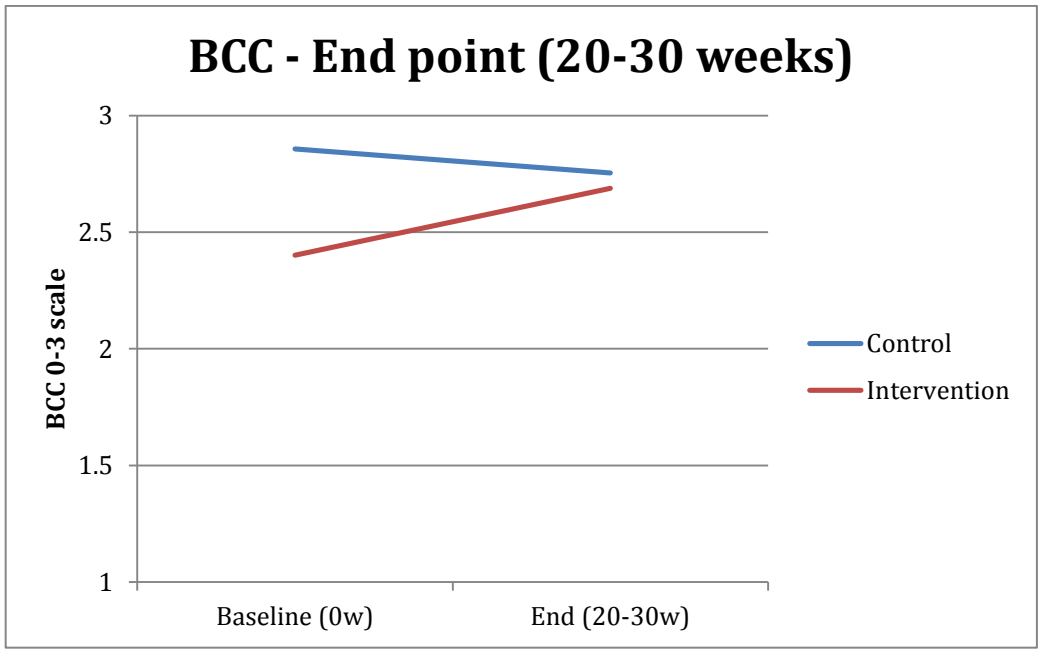


Figure 3 – Mean BCC scores for intervention and control groups at intervention baseline (0 weeks) and end-point (20-30 weeks)



In the absence of being able to test the parallel trends assumption we use a conditional parallel trends assumption and hence control for a range of sociodemographic and health factors in the DiD analysis in equation (1) (see variable X_i). The trends in Figures

2 and 3 suggest that there may be an impact of RADIQL on these outcomes and this will be tested more rigorously in the DiD analysis.

We perform two rounds of DiD analysis:

i) Where y_{i1} is set at the intervention end point (20-30weeks) and y_{i0} is set at the baseline period (0 weeks).
ii) Where y_{i1} is set at the intervention follow up point (40+ weeks) and y_{i0} is set at the baseline period (0 weeks). This model will pick up the longer term impacts of RADIQL.

These models are run separately for the ME and BCC outcomes and hence we run four DiD models in total.

4. Results

4.1. Descriptive statistics

Table 2 - Descriptive statistics

	Overall		Control		Intervention	
	n	mean	n	mean	n	mean
ME_Baseline (0-10)	48	6.92	23	7.38	25	6.49
ME_End-point (0-10)	48	7.19	23	7.14	25	7.25
ME_Follow-up (0-10)	45	6.72	23	6.91	22	6.52
BCC_Baseline (0-3)	48	2.62	23	2.86	25	2.40
BCC_End-point (0-3)	48	2.72	23	2.75	25	2.69
BCC_Follow-up (0-3)	39	2.67	22	2.86	22	2.49
Age (years)	45	86.28	20	87.18	25	85.56
Female	45	0.82	20	0.85	25	0.80
White	40	0.70	18	0.83	22	0.59
Dementia type: AD	43	0.33	18	0.17	25	0.44
Dementia type: Other	43	0.09	18	0.11	25	0.08
Dementia type: Mixed	43	0.07	18	0.06	25	0.08
Dementia type: Other	43	0.511	18	0.67	25	0.40
Date of admission (years)	43	2.69	18	3.22	25	2.31
Communication problem	40	0.25	17	0.12	23	0.35
Mobility problem	44	0.64	19	0.74	25	0.56

Note: Sample restricted to respondents for whom we have baseline (0w) and end-point (20-30w) data.

4.2. Impact analysis

4.2.1. Difference in outcomes between baseline and end-points/follow-up

Paired t-tests showed that there was a statistically significant positive change between ME levels at baseline and the end-point (20-30 weeks) for the treatment group (p=0.001). The change in ME levels between baseline and follow up (40+ weeks) was not statistically significant (p= 0.828).

Paired t-tests also showed that there was a statistically significant positive change between BCC levels at baseline and the end-point (20-30 weeks) for the treatment group (p=0.043). The difference in BCC levels between baseline and follow up (40+ weeks) was not statistically significant (p 0.914).

This simple analysis suggests that the impact of RADIQL (if any) is likely to be observed more over the short term. In the next stage of analysis we compare these trends with trends in the control group in DiD analysis.

4.2.2. Difference in difference analysis

4.2.2.1. ME at baseline (0 weeks) and end point (20-30weeks)

We find a significant positive association between the RADIQL intervention and ME at end point (20-30 weeks) after controlling for a range of background factors at baseline. Since the ME score is measured in 2-point intervals the coefficient on 'Treatment' needs to be doubled to show the effect on ME. **The effect of the RADIQL treatment at the end point (20-30 weeks) is a 2.74 point increase in the ME score (CI (95%): 0.68, 4.79)**. This effect is **significant at the 5% level**. To put this impact into perspective, a difference in ME of 2.74 amounts to an improvement of around 42% in patients' quality of life in the intervention group.

Table 4 – DiD results for ME score between baseline (0 weeks) and end point (20-30 weeks)

	Coefficient	Standard error
Treatment (RADIQL)	1.368**	0.495
Age	-0.019	0.031
Female	0.261	0.536
White	0.078	0.502
Dementia type: AD	-0.182	0.473
Dementia type: VD	0.45	0.471
Dementia type: Mixed	1.615**	0.693
Date of admission	0.033	0.098
Communication problem	-0.056	0.43
Mobility problem	0.408	0.434
Constant	0.511	2.642
Observations	32	
r2	0.463	

*Legend: * p<0.1; ** p<0.05; *** p<0.01; Reference group: (i) for Treatment, ref = control group; (ii) for gender ref = male; (iii) for white ref = non-white; (iv) for dementia type ref = other form of dementia; (v) for communication problem ref = fluent; (vi) for mobility problem ref = walking unassisted. Heteroskedasticity-robust standard errors.*

4.2.2.2. ME at baseline (0 weeks) and follow-up point (40+ weeks)

The RADIQL intervention had a positive but statistically insignificant effect on ME scores at the follow-up point (40+ weeks). This suggests that the impact of RADIQL may be more pronounced in the short term although we have to acknowledge two important caveats here. First, given that the impact is smaller over the longer term it might be that a larger sample size would help to show impacts over the longer term (40+ weeks) as well. Second, the 40+ week analysis is based on an extrapolated figure for the control group, which as we say above may overstate the ME and BCC levels in the control group at the 40-week point. This would bias the results and show an insignificant impact where in fact there may have been one. We, therefore, have greater confidence in the analysis conducted at the end point (20-30 weeks) (see section 2.2.2.1.).

Table 5 – DiD results for ME score between baseline (0 weeks) and follow-up point (40+ weeks)

	Coefficient	Standard error
Treatment (RADIQL)	1.004	0.848
Age	-0.025	0.056
Female	-0.376	0.784
White	1.37	0.892
Dementia type: AD	-0.425	0.717
Dementia type: VD	0.048	1.219
Dementia type: Mixed	1.299	1.333
Date of admission	-0.122	0.082
Communication problem	-0.402	0.706
Mobility problem	-0.106	0.673
Constant	1.196	5.109
Observations	30	
r2	0.304	

*Legend: * p<0.1; ** p<0.05; *** p<0.01; Reference group: (i) for Treatment, ref = control group; (ii) for gender ref = male; (iii) for white ref = non-white; (iv) for dementia type ref = other form of dementia; (v) for communication problem ref = fluent; (vi) for mobility problem ref = walking unassisted. Heteroskedasticity-robust standard errors.*

4.2.2.3. BCC at baseline (0 weeks) and end point (20-30weeks)

We find a significant positive association between the RADIQL intervention and BCC at end point (20-30 weeks) after controlling for a range of background factors at baseline. **The effect of the RADIQL treatment at the end point (20-30 weeks) is 0.6 point increase in the BCC score on a scale of 0-3 (CI (95%): 0.004, 1.188).** This effect is **significant at the 5% level.** A difference in BCC of around 0.6 points on a 0-3 scale amounts to an improvement of around 25% in patients' positive behaviour in the intervention group.

Table 6 – DiD results for BCC score between baseline (0 weeks) and end point (20-30 weeks)

	Coefficient	Standard error
Treatment (RADIQL)	0.596**	0.285
Age	0.003	0.022
Female	0.5	0.465
White	0.161	0.443
Dementia type: AD	0.091	0.285
Dementia type: VD	0.758	0.73
Dementia type: Mixed	0.746*	0.428
Date of admission	0.089	0.055
Communication problem	0.231	0.316
Mobility problem	-0.038	0.281

Constant	-1.426	1.874
Observations	32	
r2	0.332	

*Legend: * p<0.1; ** p<0.05; *** p<0.01; Reference group: (i) for Treatment, ref = control group; (ii) for gender ref = male; (iii) for white ref = non-white; (iv) for dementia type ref = other form of dementia; (v) for communication problem ref = fluent; (vi) for mobility problem ref = walking unassisted. Heteroskedasticity-robust standard errors.*

4.2.2.4. BCC at baseline (0 weeks) and follow-up point (40+ weeks)

The RADIQL intervention had a positive but statistically insignificant effect on BCC scores at the follow-up point (40+ weeks). This suggests that the impact of RADIQL may be more pronounced in the short term although we have to acknowledge the two important caveats discussed above about sample size and extrapolation of figures. As with the ME analysis we, therefore, have greater confidence in the analysis conducted at the end point (20-30 weeks) (see section 2.2.2.3.).

Table 7 – DiD results for BCC scores between baseline (0 weeks) and follow-up point (40+ weeks)

	Coefficient	Standard error
Treatment (RADIQL)	0.06	0.357
Age	-0.015	0.025
Female	0.81	0.486
White	-0.305	0.492
Dementia type: AD	0.025	0.322
Dementia type: VD	0.125	0.631
Dementia type: Mixed	0.363	0.416
Date of admission	0.038	0.061
Communication problem	0.295	0.261
Mobility problem	-0.232	0.292
Constant	0.866	1.992
Observations	29	
r2	0.194	

*Legend: * p<0.1; ** p<0.05; *** p<0.01; Reference group: (i) for Treatment, ref = control group; (ii) for gender ref = male; (iii) for white ref = non-white; (iv) for dementia type ref = other form of dementia; (v) for communication problem ref = fluent; (vi) for mobility problem ref = walking unassisted. Heteroskedasticity-robust standard errors.*

4.3. Summary of results

We have used a number of methods that all demonstrate that the RADIQL programme is associated with improvements in quality of life (ME) and positive behaviour (BCC) at the end point (20-30 weeks) despite small sample sizes.

There are statistically significant increases in ME and BCC scores from baseline in the RADIQL treatment group. The **DiD analysis** demonstrates that the RADIQL programme is associated with higher ME and BCC scores at the end point period (20-30 weeks). The RADIQL treatment group's scores increased in comparison to trends in a control group even after controlling for baseline differences across the two groups.

Our preferred results are for the DiD analysis for baseline (0 weeks) to end-point (20-30 weeks) and these results are further backed up by two other aspects of evidence. First, we can use the trends of the DCM score (an aggregate of ME and BCC scores) in Figure 1 to effectively show the results of an **interrupted times-series study** (where treatment is administered on numerous occasions). Interrupted time-series studies provide strong evidence of causal effect when outcomes systematically change with each treatment since it is very unlikely that some other confounding factor (which could also explain changes in the outcomes of interest) would be occurring at the exact same time as the treatment. The interrupted time-series trends in Figure 1 show that the RADIQL programme leads to short term improvements in the outcomes of interest. The DiD analysis shows that these improvements can be traced over the period up to the end point (20-30 weeks).

The second element of supporting evidence comes from the pre-post trends in ME scores for the treatment and control groups shown in Figure 2. This observed pattern of trends is "particularly amenable to causal interpretation" (Shadish et al., 2002, p.143). The '**cross-over**' in outcome trends observed for ME generally rules out biases such as maturation effects and regression to the mean effects allowing us to be more confident of a causal relationship. Although a cross-over is not observed for the BCC scores there is close to being one and we may find one with reliable observations after the end-point period (rather than extrapolated data points). However, still the observed patterns for BCC in Figure 3, whereby the treatment group was performing worse at baseline but 'catch up' to some extent with the control group - although not as convincing as the cross-over observed for ME scores in Figure 2 - are often interpreted as providing evidence of a causal relationship, although issues such as maturation effects may need to be considered (Keller and Holland, 1981; Shadish et al., 2002).

We do not find any impact of RADIQL on the outcomes of interest at the follow-up point (40+ weeks). This may be because the effects of RADIQL are more short term, although we must acknowledge that the lack of a statistically significant effect may be due to a sample size issue or because our extrapolated outcomes for the control group at week 40+ are incorrect. In light of these issues we have greater confidence in the shorter term analysis over the 20-30 week period.

5. Cost analysis of the RADIQL programme

In order to perform a form of CEA we gathered data on the costs of running the RADIQL program at the central level (costs incurred by Age Exchange in managing the project) and for each of the care homes at which the RADIQL program was implemented. These costs represent **additional costs** incurred in running the 24-week RADIQL program, over and above the business as usual costs of running a care home.

Centralised costs included staff costs (contracted costs invoiced by Age Exchange staff and consultant workers), staff training, materials and facilities.

Care home staff costs are calculated using an average daily wage of £90 for care staff workers. Care home management wages range between £19,032 - £37,375 pa. Non-management wages range widely, from nurses £18,648 - £33,587 pa to £12,504 - £17,360 pa for care workers[5]. Since we were unable to distinguish between hierarchical wage structures in our cost data we applied an average wage of £90 per day to all care home staff costs.

Refreshments were provided at each RADIQL session in the form of tea, juice and biscuits. We estimated the cost of refreshments for an average of 8 people at £20 per session, multiplied across 8 sessions to £160 per care home for the 24-week period. We estimated an opportunity cost for facilities used in the RADIQL intervention. RADIQL was delivered in care home sitting rooms (n=3), dining rooms (n=2), and in one case an activity room. We used a proxy price of the cost to RADIQL of renting out a similar room for the length of the RADIQL intervention (four hours) at £25 per hour and multiplied this by the number of hours they were occupied by the RADIQL intervention (48 hours in total).

Other care home costs included training costs, where relevant. We did not include any extra costs for materials at the care home level since most care homes provided their own equipment, notably musical instruments and stereos for performing the RADIQL intervention (at three sites the artists provided their own instruments). We have therefore assumed that this equipment does not hold an opportunity cost, which we believe is a plausible assumption.

We outline the full costs associated with the RADIQL program in Table 8. The first row shows costs incurred by Age Exchange: staff cost, training, materials and overheads. The total cost to Age Exchange of running the 24-week RADIQL project was estimated at £101,158.

The remaining rows show care home hours, staff costs, training, facilities and refreshment costs. We estimate total cost of running the RADIQL intervention across the six intervention care home sites as £18,870. This is an average cost of £3,145 per care home.

[5] Source: http://www.payscale.com

In the final cell we calculate a total cost associated with the 24-week RADIQL project of £120,028. We note that these costs may overstate the cost of running just the 24-week programme since once staff have undertaken the training they can run the RADIQL interventions on numerous occasions. The training costs are therefore not just costs associated with the 24-week programme. The costings figures when used in CEA will, therefore, produce conservative estimates of programme effectiveness.

Table 8 – Costs associated with the 24-week RADIQL programme

	Staff days (n)	Staff costs (£)	Training (£)	Materials (£)	Facilities (£)	Refresh-ments (£)	Total
Age Exchange	516	55,200	19,710	4,490	21,758	-	£101,158
Dulwich Care Home (A)	24.5	2,205	0	0	1,200	160	£3,565
Windmill Lodge (B)	24.5	2,205	0	0	1,200	160	£3,565
Bluegrove (C)	13	1,170	675	0	1,200	160	£3,205
Aashna Care Home (D)	12.5	1,125	0	0	1,200	160	£2,485
Minnie Kidd(E)	12.5	1,125	0	0	1,200	160	£2,485
Ann Moss (F)	24.5	2,205	0	0	1,200	160	£3,565
Total	627.5	65,235	20,385	4,490	28,958	960	**£120,028**

6. Cost-effectiveness analysis of the RADIQL programme

CEA is a recommended and internationally-accepted form of evaluation which is commonly used in health assessments where the valuation of outcomes is notoriously difficult. A CEA study can be used to rank alternative policies with the same objectives and measured outcomes.

CEA is most meaningful as a measure of programme effectiveness when all costs related to the programme have been calculated and when the outcome measure used in CEA captures most or all of the benefits associated with the programme. In such scenarios CEA will often replicate the findings we would get from a full CBA in terms of the ranking of the effectiveness of different projects.

In the evaluation of RADIQL we have accounted for all of the main costs associated with the programme. The remaining question is therefore whether the outcome measures (ME and BCC) capture the benefits of the programme. Since two outcomes were measured it would suggest that one of the outcomes on its own is not a sufficient measure of the programme's benefits. There may also be other benefits to the programme that are not captured by either the BCC or ME measure, such as the impacts and benefits for members of the patient's family. It can probably be argued that the CEA analysis, therefore, is likely to understate the benefits of the RADIQL programme and that full CBA study may show the programme to be more effective. However, as discussed above CEA is the only viable option to undertake here given the data at our disposal.

It is standard to measure the cost-effectiveness of an intervention *incrementally* in comparison to the status quo or another programme using the following formula (Boardman et al., 2010):

$$CER_{ij} = \frac{(C_i - C_j)}{(E_i - E_j)} \qquad (2)$$

Where CER_{ij} = the cost-effectiveness ratio (CER) of programme i relative to programme j; E is the effectiveness of a programme in terms of impacts on an outcome; and C are the costs of the programme, measured where possible as opportunity costs.

As discussed above, in the evaluation of RADIQL all costs and effectiveness figures relate to the *additional* amounts relative to status quo activities in care homes. In this case the CER ratio in equation (2) simplifies to:

$$CER_{ij} = \frac{C_{(additional)}}{E_{(additional)}} \qquad (3)$$

The results from Table 8 show that $C_{(additional)}$ = £120,028

The impact figures are calculated for the number of people that participated in the programme. It should represent the impact on the participants associated with the

overall costs of running the programme (£120,028). Thirty five people completed the RADIQL programme.

The results from the DiD analysis show the average (individual level) effect of the RADIQL programme on ME and BCC [6]. Using the results for the 20-30 week impact of RADIQL (since as explained above they are more robust) the total impacts are estimated as follows:

$E_{(additional)}$ (ME) = 2.74 x 35 = 95.9

$E_{(additional)}$ (BCC) = 0.596 x 35 = 20.86

Table 9 – Cost-effectiveness analysis of the RADIQL programme (baseline to end point)

Outcome	Additional costs of RADIQL	Additional impact of RADIQL (per person)	Number of participants	Total additional impact of RADIQL	CER (cost per unit improvement in outcome)
Mood/engagement (ME)	£120,028	2.74	35	95.9	£1,252
Behaviour category codes (BCC)	£120,028	0.60	35	20.86	£5,754

Notes: All figures are estimated as additional to the status quo (current care given in care homes).

The CER shows the cost of creating an additional <u>one point change</u> in the outcome score.

- **It costs on average £1,252 to increase ME by one unit; and**
- **It costs on average £5,754 to increase BCC by one unit**.

This could be a one unit change for one person or a one unit change spread over numerous people. Since the ME and BCC scores are measured on different scales the two CER figures cannot be compared.

The CER figures relate to changes in comparison to business as usual; the CER compares the additional costs of running RADIQL against the additional impacts of RADIQL all in comparison to current (business as usual) care in care homes. The impacts and CER results must be interpreted within the context of the current study. They relate to the effectiveness of the RADIQL programme that was delivered in the particular way that it was in the care homes in the study. If RADIQL were implemented in different care homes, with different patient groups, over different lengths of time, and at a different scale then the cost-effectiveness of the programme may differ.

To put the results into perspective a one unit change in the ME score represents a 13% change from the average scores at baseline for the intervention group. And a one unit change in the BCC score represents a 35% change from the average scores at baseline for the intervention group.

[6] The impact estimate from the DiD analysis is the average treatment effect on the treated (ATT).

By themselves these ratios *do not* indicate whether the programme is efficient or 'worthwhile' overall. To make this assessment we require a point of comparison. This could be a comparison against other intervention(s) where the intervention with the *lowest* CER figure would be the best. Or the comparison could be against a given threshold that is deemed acceptable. If the CER figure for RADIQL is *lower* than the accepted threshold (eg, this could be a threshold of £2,000 per unit improvement in the ME outcome) then the programme is effective and worthwhile.

We currently do not have estimates of programme effectiveness (measured through CER) for other dementia-related interventions for the same outcome variables. For now then, decision makers will need to make a judgement on the CER results for RADIQL in Table 9. This should be based on a judgement of whether the CER results for ME and BCC are acceptable which may be based on, for example, a reference to current levels of expenditure in care homes for people with dementia (ie, are costs of £1,252 and £5,754 relatively small in comparison to the general expenditure on care homes?). There may also be other ways that a judgement about the CER results could be made in absence of a direct comparison programme.

The other approach would be to apply monetary values to the ME and BCC impacts so that CBA can be performed. This would then make a straight assessment between benefits and costs possible without reference to the performance of other programmes since the costs and benefits of RADIQL would be calculated in the same metrics. As discussed, it is not possible to value ME and BCC outcomes using current research and data but this should be an area for future research to further develop the evaluation methodology of RADIQL.

7. Conclusion and discussion

The aim of this study was to provide a full assessment of the effectiveness of the RADIQL programme using HM Treasury Green Book consistent methods. The RADIQL intervention was tested on six care homes in the London boroughs of Lambeth and Southwark, and the results compared against six care homes designated as control sites. Thirty five people completed the 24-week RADIQL programme. Individuals in the control group received no change to their normal care regime.

Patients' response to the RADIQL intervention was measured using the Dementia Care Mapping (DCM) quality of life measure. The two main outcomes that have been measured in this paper are positive behaviour measures (measured through Behaviour Category Codes (BCC)) and a quality of life (QoL) measure (measured as a subjective mood and engagement (ME) score). Data on BCC and ME were captured at two points - at baseline and at an end-point range of 20-30 weeks - for both the intervention and control groups. Follow-up observations were performed on the intervention group only at around 42 weeks after the baseline period.

We used difference in difference (DiD) statistical analysis to assess the impact of RADIQL on BCC and ME scores. The strength of this method comes from the fact that the approach exploits variation over two dimensions: changes over time and changes across groups in the outcome. We find a significant positive association between the RADIQL intervention and ME and BCC scores at the end point. The average effect of the RADIQL treatment at the end point is a 2.74 point increase in the ME score. This amounts to an improvement of around 42% in patients' quality of life in the intervention group. The average effect of the RADIQL treatment at the end point on BCC is a 0.6 point increase in the BCC score on a scale of 0-3. This amounts to an improvement of around 25% in patients' positive behaviour in the intervention group.

We performed cost-effectiveness analysis (CEA) to calculate the cost of a unit improvement in ME and BCC levels associated with the RADIQL intervention. CEA is a Green Book recommended form of evaluation in the absence of full cost-benefit analysis (CBA). We compared the additional costs of running RADIQL against the additional ME and BCC impacts in comparison to current (business as usual) care in care homes. We found that it costs on average £1,252 to increase ME by one unit; and £5,754 to increase BCC by one unit.

To put the results into perspective a one unit change in the ME score represents a 13% change from the average scores at baseline for the intervention group, while a one unit change in the BCC score represents a 35% change from the average scores at baseline for the intervention group. We note that this could be a one unit change for one person or a one unit change spread over numerous people.

In order to make an assessment of whether the RADIQL programme is efficient overall we require a point of comparison, for instance another intervention, or an objective threshold that is deemed acceptable. Alternatively, we could apply monetary values to the ME and BCC impacts so that full CBA can be performed. This would then make possible a straight assessment between the benefits and costs of RADIQL without

reference to the performance of other programmes. It is not possible to value ME and BCC outcomes using current data, and so valuation of these outcomes is a key area for future research to further develop the evaluation methodology of RADIQL.

References

Allen, J. (2008). Older People and Wellbeing. Institute for Public Policy Research.

Boardman, A., Greenberg, D., Vining, A., and Weimer, D. (2010). Cost-Benefit Analysis: Concepts and Practice. Prentice Hall.

Keller, R. and Holland, W. (1981). Job change: A naturally-occurring field experiment. Human Relations, 134, 1053-1067.

Khuri, S. (2010). Information, Technology in Bio- and Medical Informatics, ITBAM 2010: First International Conference, Bilbao, Spain, September 1-2, 2010, Proceedings. Springer Science & Business Media.

NICE. (2014). Improving the Lives of People in Care Homes. London, UK: The National Institute for Health and Care Excellence. https://www.nice.org.uk/news/feature/improving-the-mental-wellbeing-of-older-people-in-care-homes

Shadish, W., Cook, T., Campbell, D. (2002). Experimental and quasi-experimental designs for generalised causal inference. Houghton Mifflin Company. New York.

Woods, B., Spector, A., Jones, C., Orrell, M., & Davies, S. (2005). Reminiscence therapy for dementia. Cochrane Database Systematic Review, 2.

Zeilig, H., Killick, J., & Fox, C. (2014). Mark Making: a critical review of the value of arts and culture for people living with a dementia. University of the Arts: London: Arts and Humanities Research Council.

SIMETRICA LTD
Crowne House
72 Hammersmith Road
W14 8TH

Email: information@simetrica.co.uk
Phone: 020 7559 1344
www.simetrica.co.uk
@_simetrica

Company number: 9035322 (England and Wales)