An empirical comparison of the EQ-5D-5L, DEMQOL-U and DEMQOL-Proxy-U in a post-hospitalisation population of frail older people living in residential aged care

Ratcliffe J[1], Flint T[1], Easton T[2], Killington M[2], Cameron ID[3], Davies O[2], Whitehead C[2], Kurrle S[3], Miller M[4], Liu E[2], Crotty M[2]

[1] Flinders Health Economics Group, School of Medicine, Flinders University

[2] Department of Rehabilitation and Aged Care, School of Health Sciences, Flinders University

[3] John Walsh Centre for Rehabilitation Research, Kolling Institute, University of Sydney

[4] Department of Nutrition and Dietetics, School of Health Sciences, Flinders University

Running head: quality of life of frail older people in residential aged care

Word count: 4138 (excluding references)

Compliance with Ethical Standards

This study was approved by the Southern Adelaide Clinical Human Research Ethics Committee (SAC HREC EC00188) (Project no:20.12). All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. Informed consent was obtained from all individual participants included in the study.

Acknowledgements

The authors gratefully acknowledge funding provided by the National Health and Medical Research Council (NHMRC) Partnership Centre on Dealing with Cognitive and Related Functional Decline in Older People, NHMRC Project ID GNT9100000.

Key points:

- In contrast to the proliferation of studies reporting upon an assessment of the quality of life of older people in community based setting there is a paucity of evidence relating to the health related quality of life of older people living in residential care

- This paper provides an empirical assessment of the measurement properties of the newly developed DEMQOL-U and DEMQOL Proxy-U instruments to the EQ-5D-5L and its proxy version in a population of frail older people residing in residential aged care in the post-acute period following a hip fracture.

- The findings highlight the central importance of self-versus proxy assessment and the choice of preference based instrument for the measurement and valuation of health related quality of life in older people exhibiting cognitive decline, dementia and other co-morbidities.

Abstract

Purpose: To empirically compare the measurement properties of the DEMQOL-U and DEMQOL Proxy-U instruments to the EQ-5D-5L in a population of frail older people living in residential aged care in the post-hospitalisation period following a hip fracture.

Methods: A battery of instruments to measure health related quality of life (HRQoL), cognition, and clinical indicators of depression, pain and functioning were administered at baseline and repeated at 4 weeks follow up. Descriptive summary statistics were produced and psychometric analyses were conducted to assess the levels of agreement, convergent validity and known group validity between clinical indicators and HRQoL measures.

Results: There was a large divergence in mean (SD) utility scores at baseline for the EQ-5D-5L and DEMQOL-U [EQ-5D-5L mean 0.21 (0.19); DEMQOL-U mean 0.79 (0.14)]. However, at 4 weeks a stronger degree of convergence was evident [EQ-5D-5L mean 0.45(0.38); DEMQOL-U mean 0.58 (0.38). The EQ-5D was more responsive to the physical recovery trajectory experienced by frail older people following surgery to repair a fractured hip, whereas the DEMQOL-U and DEMQOL-Proxy-U instruments were more responsive to the changes in cognitive functioning often experienced by frail older people in this period.

Conclusions: This study presents important insights into the HRQoL of a relatively under-researched population of post- hospitalisation frail older people in residential care. Further research should investigate the implications for economic evaluation of self-versus proxy assessment of HRQoL and the choice of preference based instrument for the measurement and valuation of HRQoL in older people exhibiting cognitive decline, dementia and other co-morbidities.

1. Introduction

A hip fracture represents a sentinel event in the life of an older person with estimates indicating that 25-35% will die within a year of incurring a fractured hip [1] whilst only 40% will rediscover the level of mobility they enjoyed prior to the fracture [2]. In Australia, it has been estimated that approximately 30% of all hip fractures occur in residential aged care facilities [3]. In contrast to the proliferation of studies reporting upon the health related quality of life of older people recovering from hip fracture in community based settings there is a lack of evidence relating to the health related quality of life of older people recovering from hip fracture living in residential care. Those studies which have been conducted in this setting have tended to focus more generically on the measurement of quality of life of residents and have typically reported quality of life to very poor in this population [4-6].

The DEMQOL and DEMQOL-Proxy were developed as condition specific instruments to capture the measurement of HRQoL of individuals with cognitive decline and dementia where the DEMQOL was designed to be self-completed by the individual and the DEMQOL-Proxy was designed to be completed by a suitable proxy e.g. close family member or a carer [7,8]. In their original form both the DEMQOL and the DEMQOL-Proxy are not suitable for use in economic evaluation as they provide summary scores that are not preference based. However, the recent development of the DEMQOL-U and DEMQOL-Proxy-U using general population preference values has facilitated the application of these measures in economic evaluations [9,10].

The EQ-5D represents the world's most widely used preference based measure of HRQoL and the instrument is well known for its reliability, responsiveness and validity [11]. The

acceptability and feasibility of the EQ-5D for administration with individuals exhibiting mild to moderate cognitive impairment and living in residential aged care and has been demonstrated in a number of published studies internationally [12-15]. For individuals who are unable to self-complete the EQ-5D due to more severe levels of cognitive impairment a proxy version is available (CEQ-5D). A recent study analysed the psychometric properties of the EQ-5D rated by proxy and demonstrated its validity and reliability in individuals living in residential care with different stages of dementia [15]. The study authors concluded that proxy assessment may improve the feasibility of the EQ-5D for individuals at advanced stages of dementia who are unable to self-complete the instrument.

The high prevalence of hip fractures occurring in residential aged care facilities highlight the importance of providing cost-effective post-hospitalisation treatment and care pathways [6]. Within economic evaluation, benefits are most often captured by quality adjusted life years (QALYs) through the administration of preference based measures such as the DEMQOL-U, DEMQOL-Proxy-U and EQ-5D [11]. Selecting the most appropriate instrument/s for measuring and valuing health related quality of life (HRQoL) in frail older people recovering from hip fracture and living in residential care is not straightforward. Hip fracture patients returning to residential care from hospital typically experience differential care and rehabilitation treatment pathways that may impact upon their physical recovery trajectory [2]. They often also experience delirium and consequently are at greater risk of experiencing accelerations in cognitive decline and dementia [2,6]. Currently there is a paucity of empirical evidence to guide the selection of the most appropriate outcome measure/s for application in this population and setting [6,11].

The main aim of this study was to conduct an empirical assessment of the measurement properties of the newly developed DEMQOL-U and DEMQOL Proxy-U instruments to the EQ-5D-5L and its proxy version in a population of frail older people residing in residential aged care in the post-hospitalisation period of recovery (0-4 weeks) following surgery to repair a hip fracture. It was hypothesized *a priori* that although the DEMQOL-U, DEMQOL Proxy-U and EQ-5D-5L instruments were all designed to measure the same concept of utility on an equivalent quality adjusted life years (QALY) scale (where zero represents the state dead and one represents the state of full health), the utilities associated with each of these instruments and the magnitude of change over time in utilities may differ. In particular it was hypothesized that the EQ-5D would be more responsive to the physical recovery trajectories experienced by frail older people recovering from a hip fracture, where as it was hypothesized that the DEMQOL-U and DEMQOL-Proxy-U instruments would be more responsive to the changes in cognitive functioning and symptoms of dementia often experienced by frail older people living in residential care in the period following a hip fracture.

2. Methods

Data source and Measures

The data utilised for this study emanates from a randomised controlled trial to investigate the clinical and cost effectiveness of a multidisciplinary rehabilitation service to hip fracture patients living in residential aged care [6]. A battery of instruments to measure quality of life, cognition, and clinical indicators of depression, pain and functioning were administered at baseline and repeated at 4 weeks follow up with individuals living in residential care and recovering from surgery to repair a fractured hip.

In accordance with the recommendations produced by the instrument developers both the DEMQOL-U and DEMQOL-Proxy-U were administered in individuals with mild to moderate dementia, whereas for individuals with severe dementia, only the DEMQOL-Proxy was administered [8]. For proxy assessment, family carers were utilised wherever possible. For a small number of cases where a family carer was unavailable, proxy assessment was undertaken by a residential care staff member involved in the day to day care of the individual. The scoring algorithms pertaining to the DEMQOL-U and DEMQOL-Proxy-U classification systems were derived using the time trade off elicitation technique in a UK general population sample. The resulting utility scores lie on the zero to one quality adjusted life years (QALY) scale where zero represents the state dead and one represents the state of full health. The utility scores for the DEMQOL-U range from 0.243 to 0.986 and the DEMQOL-Proxy-U from 0.363 to 0.937 [9,10].

The EQ-5D is a commonly used preference based measure of health status designed for completion by patients and/or members of the general population, however it can also be completed by a proxy (CEQ-5D). The EQ-5D-5L was recently developed in order to improve sensitivity and reduce ceiling effects relative to the original 3 level version [16]. Although no guidance currently exists from the instrument developers in relation to self-complete versus proxy administration of the EQ-5D in individuals with cognitive impairment and dementia, evidence from a study recently reported upon in this journal indicates that people with mild to moderate dementia are able to rate their own HRQoL using the EQ-5D and proxy ratings were found to be consistently different from individual ratings [17]. For the purposes of the randomised controlled trial from which this study emanates, clinical judgement was used to

reflect upon the individual's ability to understand and tolerate the task. Self-completion of the EQ-5D-5L was encouraged wherever possible on the basis that the individual is normally viewed as the best judge of their own HrQoL [11]. When engagement with the task was reduced to such a level that the clinician viewed it as unreasonable to commence or continue, then proxy assessment via a family member (or for a small number of cases a staff member in the absence of a family carer) was utilised. A scoring algorithm for the EQ-5D-5L is available based upon the time trade off and discrete choice experiment approaches with a UK general population sample. The resulting utility scores range from -0.281 to 1 (where states with a score less than zero are considered worse than being dead) [18].

Cognitive functioning was assessed by the Mini-Mental State Examination (MMSE), a routinely administered brief instrument for the measurement of global cognitive function [19]. The Mini Mental State Examination (MMSE) is the most commonly used test for complaints of problems with memory or other mental abilities internationally. It is routinely used by clinicians to help diagnose dementia and to help assess its progression and severity [20]. MMSE scores below 10 are indicative of severe dementia, scores between 10 and 20 suggests moderate dementia and finally scores greater than 20 suggest mild dementia to non-existent cognitive impairment [21]. The Cornell Scale for Depression in Dementia (CSDD) is a 19-item instrument completed by clinical assessment that screens for signs of depression specifically among dementia patients [22, 23]. Scores >11 indicate probable depression while scores >18 indicate definite depression.

Individual's functional ability or independence was measured by the Modified Barthel Index (MBI) [24]. The MBI is a 10-item instrument that produces a score between 1 and 100 with a

score of 0-20 indicating total dependence; 21-60 indicating severe dependence; 61-90 indicating moderate dependence; 91-99 indicating slight dependence; and finally a score of 100 indicates that the individual is fully independent in basic daily activities such as showering, dressing and mobility. Pain levels were assessed using the Pain Assessment in Advanced Dementia Scale (PainAd) [25], an observational instrument designed for administration by a clinician to measure levels of pain in patients with advanced dementia. The instrument produces scores between 0 and 10 with 0 being no pain and 10 indicating severe pain.

Data Analysis

Descriptive Analysis

Descriptive summary statistics were produced initially to provide an overview of the study population in terms of their socio-demographic characteristics and clinical indicators of depression, pain and functioning at baseline. Summary statistics of the quality of life measures of interest, the DEMQOL-U and DEMQOL-Proxy-U and the EQ-5D-5L and CEQ-5D-5L, were generated to show how individuals had responded to these instruments at each time point.

Agreement

In order to assess the levels of agreement between the different measures of quality of life, Bland-Altman plots were created. Bland-Altman plots were generated to explore agreement between DEMQOL-U and EQ-5D-5L and also DEMQOL-Proxy-U and CEQ-5D-5L both at baseline and at the week 4 follow up period to assess the extent to which similar responses were given to each instrument. As the health states were reported either by the individuals themselves or their carer the patient's 'real' health state is an unknown. Therefore, the x axis of the Bland-Altman plots uses the average of the utilities values derived from the two instruments in

question. The y axis consists of the difference between the two utility scores. For example for the Bland-Altman plot assessing agreement between DEMQOL-U and EQ-5D-5L, x = (DEMQOL-U score + EQ-5D-5L score)/2 and y= DEMQOL-U score – EQ-5D-5L score.

Individual and Proxy agreement

In addition to exploring the agreement between the different quality of life measures, agreement between self-report and proxy report for the same measure was analysed. To show the level of agreement graphically Bland-Altman plots were constructed in a similar way to those for agreement between the different instruments. As we do not know the 'true' health state of the individual the x axis is the average of the self-reported and proxy reported utility values while the y axis is the difference between the two utility values. This analysis was completed for the proportion of the total sample (32% at baseline and 49% at week 4) where both the DEMQOL-U and DEMQOL-Proxy-U were available for individuals. As indicated previously, the EQ-5D-5L and the CEQ-5D-5L were completed by either the individual or a proxy family member respectively and not simultaneously, therefore it was not possible to conduct this analysis for the EQ-5D.

Convergent Validity

The association between DEMQOL-U, EQ-5D-5L and their respective proxy versions, cognition, and clinical indicators of depression, pain and functioning was explored to assess the degree of convergent validity. This involved determining whether there was a strong correlation between the utility scores derived from the DEMQOL-U and EQ-5D-5L (and proxy versions) and scores relating to cognition, clinical indicators of depression, pain and functioning (MMSE,

CSDD, MBI and PainAd) using spearman correlation coefficients. Correlations greater than 0.7 were considered strong and correlations between 0.4 and 0.6 were considered moderate [26].

Known-group validity

Recommended threshold levels for severity in instruments such as the MMSE and the MBI were utilised to facilitate an assessment of known group validity. For example, it was expected that those with severe dementia (MMSE score <10) would have a lower quality of life and therefore lower utility values for the DEMQOL-U and EQ-5D-5L. Similarly, it was expected that those with mild dementia (MMSE score >20), or no cognitive impairment, would have higher utility values. Therefore, using these pre-defined threshold values from other instruments designed to measure specific aspects of health status, the utility values derived from DEMQOL-U and EQ-5D-5L were compared between severity groups to examine whether they behaved as would be expected.

3. Results

Sample Characteristics and Descriptive statistics

A total of 354 individuals were approached to be included in the study of whom N=240 (68%) were eligible and consented to participate. The socio-demographic characteristics of the study sample are presented in Table 1. Summary statistics for the EQ-5D, CEQ-5D, DEMQOL-U and DEMQOL-Proxy-U across the study period for the total sample and for the sub-sample of individuals who remained alive at the 4 week follow up period and completed the same version of each instrument at both time points are presented in Table 2. For the total sample, the mean utility scores at baseline for the EQ-5D-5L and DEMQOL-U were 0.21 and 0.79 respectively. In

agreement with our prior hypotheses the respective utilities and the direction of the change in utilities from baseline to week 4 differed for each instrument. As expected, the EQ-5D-5L improved in general at week 4 follow up with a mean score of 0.45 whereas the mean score for the DEMQOL-U at week 4 follow up had fallen to 0.58. Regarding the proxy versions of the instruments, the CEQ-5D-5L and the DEMQOL-Proxy-U, had mean scores of 0.23 and 0.63 respectively at baseline. The proxy instruments also followed the same pattern as their respective patient completed versions in that the mean score for the CEQ-5D-5L increased at week 4 follow up to 0.38 while the DEMQOL-Proxy-U mean score fell to 0.57. When complete case data only were analysed a similar pattern is evident for the EQ-5D-5L with a mean score of 0.25 at baseline and an improvement in general at week 4 follow up with a mean score of 0.58. For the DEMQOL-U, there is a slight reduction in mean scores at week 4 (0.80) relative to baseline (0.81). However for the DEMQOL-Proxy-U a slight improvement over time was evident overall with a mean score at baseline of 0.64 and 0.66 at week 4.

Agreement

Figure 1 presents Bland-Altman plots showing the level of agreement between the individual self-reported utility measures (EQ-5D-5L and DEMQOL-U) and their respective proxy versions (CEQ-5D-5L and DEMQOL-Proxy-U) at baseline. It is evident that there is little agreement between the two measures at baseline, with the majority of the responses clustering around the mean difference of 0.551. This finding indicates that individuals tended to exhibit systematically higher utility scores on the DEMQOL-U. This pattern does change at week 4 follow up (presented in Figure 2) as responses cluster around the y=0 line at higher levels of utility suggesting there is better agreement at the upper end of the utility scale. However, it is evident

that there is still disagreement at lower utility levels with the DEMQOL-U indicating higher scores on average.

Patient and Proxy agreement

The proxy reported measures indicate disagreement at baseline between the two measures at lower utility levels. As utility levels increase the responses move closer to y=0 however, they mainly cluster around the mean difference of 0.403 indicating that responses to DEMQOL-Proxy-U are higher than that of the CEQ-5D-5L. At week 4, similar to the patient reported measures, there is more agreement with a high proportion of responses around the y=0 line particularly at higher utility levels. At lower utility levels there is still disagreement and the cluster around the mean difference remains although the mean difference decreased slightly to 0.279.

Convergent Validity

Spearman correlations between the utility based health measures (and the non-utility based health measures are presented in Table 3. At baseline, of the individual self-reported measures, only the EQ-5D-5L is significantly associated with MMSE score. Although this is in the expected direction (EQ-5D-5L increases as MMSE increases) the correlation is below 0.4 and therefore cannot be classified as moderate or strong. The EQ-5D-5L and DEMQOL-U were significantly associated with CSDD, MBI and PainAd in the expected direction (utility increases as CSDD or PainAd decreases and utility increases as MBI increases). The majority of the correlations were moderate with the exception of the relationships between firstly, the CSDD and EQ-5D-5L and secondly the MBI and DEMQOL-U, which indicated low correlations. In

relation to the proxy reported instruments, the DEMQOL-Proxy-U was significantly associated with all of the non-utility measures and in the expected direction. The CEQ-5D-5L and DEMQOL-Proxy were both found to be significantly associated with the CSDD and the PainAd but only their correlations with CSDD could be classified within the moderate range. At week 4, although the correlations remained in the expected directions, fewer of the relationships were identified as statistically significant. In general the proxy reported instruments performed better than their self-complete counter parts in terms of their associations with the non-utility based health measures. However, in contrast to baseline, at week 4 none of the correlations were above 0.4 indicating relatively poor associations in general.

Known group Validity

Table 4a and 4b presents the mean patient and carer utility scores categorised into severity groups of the non-utility based measures of health at baseline and 4 weeks respectively. For cognitive impairment, as measured by the MMSE, at baseline the EQ-5D-5L and DEMQOL-Proxy-U behave as would be expected with lower utilities associated with more severe levels of cognitive impairment. The DEMQOL-U follows a similar pattern in that utility decreases as cognitive impairment increases from mild to moderate however utility increases again at levels of severe cognitive impairment. This finding was unexpected but may at least partly be explained by the relatively small proportion of individuals classified with severe cognitive impairment who self-completed the DEMQOL-U. At week 4 the pattern of the relationships is less clear overall, with more of a differentiation evident between utilities and the degree of cognitive impairment for the DEMQOL-U and DEMQOL-Proxy-U relative to the EQ-5D-5L and CEQ-5D-5L. For depression, measured by the CSDD, all of the instruments behave as one would expect at both

time-points with utilities decreasing as the severity of depression increases. For self-care, as measured by the MBI, all individuals were classified into the two most severe groups at baseline (severe dependence or total dependence). However, at week 4 there was more of a range with individuals classified from slight to total dependence. The two self-reported measures (EQ-5D and DEMQOL-U) and the DEMQOL-Proxy-U behaved as expected with utilities decreasing as dependency levels increased. For the CEQ-5D-5L and the DEMQOL-Proxy however, there was less of a clear pattern, although the small numbers in some categories may have affected these comparisons. Finally, all the utility measures behaved as expected in relation to the PainAd instrument at baseline, with utilities falling as reported pain levels increased. However, at 4 weeks the pattern was less clear with the CEQ-5D-5L in particular appearing to be relatively insensitive to pain levels as reported by the PainAd instrument at 4 weeks.

4. Discussion

This study represents the first study internationally to empirically compare the measurement properties of the DEMQOL-U and DEMQOL-Proxy-U to the EQ-5D-5L and CEQ-5D-5L in a post-hospitalisation population of older people living in residential care. Evidence is presented regarding the acceptability, validity and responsiveness of the DEMQOL-U and DEMQOL-Proxy-U dementia specific preference based HRQoL measurement system and the EQ-5D generic preference based measure in comparison with external indicators of dementia related health status. Overall, we found little evidence of agreement between the DEMQOL-U and EQ-5D, and CEQ-5D and DEMQOL-Proxy-U. This finding may be attributable to the dementia condition specific focus of the DEMQOL in contrast to the EQ-5D which has a particular focus

upon the physical functioning dimensions (such as mobility, self-care and the ability to perform usual activities) severely impacted in the immediate period following a hip fracture.

To our knowledge only one other study has empirically compared the DEMQOL-U and DEMQOL-Proxy-U to the EQ-5D-5L and CEQ-5D-5L [10]. In comparison with our study population, the study by Mulhern and colleagues was undertaken in a sample with a higher prevalence of individuals (88%) with mild to moderate dementia, with only 12% of individuals exhibiting severe dementia (MMSE score<10). In addition to higher levels of cognitive impairment, on average, the individuals in our study sample were also physically frail and due to the nature of their condition almost all were classified as totally dependent at baseline. The differences in study populations are reflected in baseline EQ-5D scores with a mean EQ-5D score of 0.21 in this study relative to a mean EQ-5D score of 0.68 in the UK study population. As expected, given the differential nature of the respective study populations in terms of the levels of cognitive impairment, the DEMQOL-U and DEMQOL-Proxy-U scores reported in this study were also lower, on average, than those reported by Mulhern and colleagues at both time points (baseline and week 4). However, despite differences in populations and study settings, there were some noteable similarities between the findings of the study by Mulhern and colleagues and the study reported here in terms of the agreement found between the DEMQOL-U and EQ-5D, and CEQ-5D and DEMQOL-Proxy-U with a higher level of agreement at the upper end of the utility scale with the majority of the disagreement occurring at the more severe end of the utility scale.

The evidence from this study in relation to the convergent validity of the instruments with other measures of cognition, depression, functioning and pain is mixed with moderate correlations being observed in general at baseline but poorer correlations at week 4. In relation to known group validity overall all of the instruments demonstrated a reasonable level of performance in relation to the severity classifications of the non-utility based health measures, with utilities falling in general as severity classifications increased. Some exceptions to this general response pattern were noted but these were likely largely explained by the relative small proportion of individuals in some categories.

Examination of the distribution of utility values indicated that the DEMQOL-U and its proxy version systematically produced higher utility values relative to the EQ-5D-5L and CEQ-5D-5L. This finding may be due to differences in the classification systems, which are sensitive to different dimensions of HRQoL and also due to large differences in the possible utility scale range. The EQ-5D-5L utility scores range from -0.281 to 1 whilst the utility scores range from 0.243 to 0.986 for DEMQoL-U and 0.363 to 0.937 for DEMQoL-Proxy-U.

It is evident that both the EQ-5D-5L and CEQ-5D-5L utility scores demonstrated significant positive improvements between baseline and Week 4. In contrast smaller changes were evident for DEMQOL-U or the DEMQOL-Proxy-U during this time period and often in a negative direction. In accordance with our prior hypotheses, this study therefore indicates that the EQ-5D was more responsive to the physical recovery trajectory experienced by frail older people recovering from surgery to repair a fractured hip, where as it is likely that the DEMQOL-U and DEMQOL-Proxy-U instruments were more responsive to the changes in delirium cognitive

functioning and dementia symptoms often experienced by frail older people in the immediate period following a hip fracture [2,6].

The study raises important questions about an individual's ability to provide a reliable self-assessment of their HRQoL relative to a proxy assessor. A limitation of this study in this regard is that dyad assessments of HRQoL were available for a proportion of the total sample (approximately one-third) and only for the DEMQOL-U and the DEMQOL-Proxy-U instruments. Investigation of the agreement in utility values between self and proxy rated HRQoL indicated only a poor to moderate level of agreement overall with proxy assessors tending to report lower HRQoL than individuals themselves. Similar findings were noted in a community based study recently reported upon in this journal by Orgeta and colleagues to assess the inter-rater agreement of self and family carer proxy ratings of HrQoL for people with mild to moderate dementia [17] and have also been observed in other studies [27, 28]. These findings have potentially important implications for the results of economic evaluation studies.

5. Conclusions

This study presents insights into the HRQoL of a relatively highly under-researched population of post- hospitalisation frail older people in residential care, previously evocatively described as members of the 'lost tribe' [29]. The findings highlight important questions relating to the choice of the most appropriate preference based instrument in post-hospitalisation populations of frail older people living in residential care with cognitive decline, dementia and other co-morbidities. and the decision as to from whose perspective (self versus proxy) HRQoL is assessed. Further

research should be conducted to explore the implications of these choices for the results of economic evaluation studies conducted in this setting.

References

[1] Australian Institute of Health and Welfare. The problem of osteoporotic hip fracture in Australia. Australian Government, Institute of Health and Welfare, Bulletin No. 76, March 2010, Canberra.

[2] Braithwaite RS, Col NF, Wong JB. Estimating Hip Fracture Morbidity, Mortality and Costs. J Am Geriatr Soc 2003; 51:364-370.

[3] Koval KJ, Zuckerman JD. Functional recovery after fracture of the hip. J Bone Joint Surg Am 1994; 76:751-8.

[4] Kane R, Kling K, Bershadsky B, et al. Quality of life measures for nursing home residents. J Gerontol A Biol Sci Med Sci 2003; 58A:240-8.

[5] Giles LC, Hawthorne G, Crotty M. Health-related Quality of Life among hospitalized older people awaiting residential aged care. Health Qual Life Outcomes 2009; 7:1-7.
[6] Crotty M, Ratcliffe J. If Mohammed won't come to the mountain, the mountain must go to Mohammed. Age Ageing 2011; 40:290-2.

[7] Smith SC, Lamping D, Banerjee S, et al. Measurement of health-related quality of life for people with dementia: development of a new instrument (DEMQOL) and an evaluation of current methodology. Health Technol Assess (Winch Eng) 2005; 9:1-93.

[8] Smith SC, Lamping D, Banerjee S, et al. Development of a new measure of health-related quality of life for people with dementia: DEMQOL. Psychol Med 2007; 37:737-46.
[7] Brazier J, Ratcliffe J. Measuring and valuing health benefits for economic evaluation. New York: Oxford University Press, 2007.

[9] Mulhern B, Rowen D, Brazier J, et al. Development of DEMQOL-U and DEMQOL-PROXY-U: generation of preference-based indices from DEMQOL and DEMQOL-PROXY for use in economic evaluation. Health Technol Assess (Winch Eng) 2013; 17:1-160.

[10] Rowen D, Mulhern B, Banerjee S, et al. Estimating Preference-Based Single Index Measures for Dementia Using DEMQOL and DEMQOL-Proxy. Value Health 2012; 15:346-356.

[11] Brazier J, Ratcliffe J. Salomon J, Tsuchiya A. Measuring and valuing health benefits for economic evaluation. Oxford University Press, Oxford, 2007.

[12] Aguirre E, Kang S, Hoare Z, et al. How does the EQ-5D perform when measuring quality of life in dementia against two other dementia-specific outcome measures? Qual Life Res 2015.

[13] Devine A, Diaz-Ordaz K, Taylor SJC, et al. The agreement between proxy and self-completed EQ-5D for care home residents was better for index scores than individual domains. J Clin Epidemiol 2014; 67:1035-1043.

[14] Gordon AL, Franklin M, Bradshaw L, et al. Health status of UK care home residents: a cohort study. Age Ageing 2014; 43:97-103.

[15] Diaz-Redondo A, Rodriquez-Blazquez C, Ayala A, et al. EQ-5D rated by proxy in institutionalized older adults with dementia: psychometric pros and cons. Geriatr Gerontol Int 2014; 14:346-53.

[16] Herdman M, Gudex C, Lloyd A et al. Development and preliminary testing of the new five level version of the EQ-5D (EQ-5D-5L). Quality of Life Research 2011; 20: 1727-1736.

[17] Orgeta V, Tudor Edwards R, Hounsome B, Orrell M, Woods B. The use of the EQ-5D as a measure of health related quality of life in people with dementia and their carers. Quality of Life Research 2015; 24: 315-324.

[18] Devlin N, Shah K, Feng Y, Mulhern B, Van Hout B (2016). Valuing health related quality of life: An EQ-5D-5L value set for England. HEDS Discussion Paper No 16.02, University of Sheffield.

[19] Folstein MF, Folstein SE, McHugh PR. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res 1975; 12:189-98.

[20] Sallam K, Amr M. The use of the mini-mental state examination and the clock drawing test for dementia in a tertiary hospital. J Clin Diagn Res 2013; 7: 484-8.

[21] Peneczy R, Wagenpfeil S, Kornossa K et al. Mapping scores onto stages: mini-mental state examination and clinical dementia rating. Am J Geriatr Psychiatry 2006; 14: 139-44.

[22] Alexopoulos GS, Abrams RC, Young RC, Shamoian CA. Cornell Scale for Depression in Dementia. Biol Psychiatry 1988; 23:271-84.

[23] Kurlowicz LH, Evans LK, Strumpf NE, Maislin G. A psychometric evaluation of the Cornell Scale for Depression in Dementia in a frail, nursing home population. Am J Psychiatry 2002; 10:600-8

[24] Shah S, Vanclay F, Cooper B. Improving the sensitivity of the Barthel Index for stroke rehabilitation. J Clin Epidemiol 1989; 42:703-9.

[25] Warden V, Hurley A, Volicer L. Development and Psychometric Evaluation of the Pain Assessment in Advanced Dementia (PAINAD) Scale. J Am Med Dir Assoc 2003; 4:9-15.

[26] Dancey CP, Reidy J. Statistics without maths for psychology: using SPSS for Windows. Harlow: Prentice Hall, 2004.

[27] Coucill W, Bryan S, Bentham P, et al. EQ-5D in patients with dementia: an investigation of inter-rater agreement. Med Care 2001; 39:760-71.

[28] Arons A, Van Der Wilt G, Krabbe P, et al. Quality of life in dementia: A study on proxy bias. BMC Med Res Methodol 2013; 13:1-8

[29] Stott D, Langhorne P, Knight PV. Multidisciplinary care for elderly people in the community. Lancet 2008; 371:699–700.

Table 1: Characteristics of the study sample at baseline

Characteristic	n	Mean (SD) or %
Participant profile		
Age	240	88.6 (5.6)
Female	178	74.2
Cognitive impairment		
MMSE total score	240	8.0 (7.8)
Mild (MMSE score >20)	18	7.5
Moderate (MMSE score 10-20)	82	34.2
Severe (MMSE score <10)	140	58.3
Severity of depression		
CSDD total score	238	10.3 (5.4)
Non-case (CSDD score <6)	46	19.3
Probable (CSDD score 6-10)	87	36.6
Definite (CSDD score >10)	105	44.1
Pain		
PainAd total score	240	1.1 (1.7)
No Pain (PainAd score 0)	137	57.1
Mild (PainAd score 1-3)	77	32.1
Moderate (PainAd score 4-6)	24	10
Severe (PainAd score 7-10)	2	0.8
Self-care		
MBI total score	240	9.3 (6.8)
Independence (MBI score 100)	0	0
Slight dependence (MBI score 91-99)	0	0
Moderate dependence (MBI score 61-90)	0	0
Severe dependence (MBI score 21-60)	12	5
Total dependence (MBI score 0-20)	228	95

Table 2: Summary statistics for EQ-5D, CEQ5D, DEMQOL-U and DEMQOL-Proxy-U across the study period

	Total sample		
Measure	**n**	**Mean (SD)**	**Median (IQR)**
EQ-5D-5L			
Baseline	82	0.21 (0.19)	0.27 (0.16-0.28)
Week 4	109	0.45 (0.38)	0.49 (0.00-0.80)
DEMQOL-U			
Baseline	71	0.79 (0.14)	0.85 (0.68-0.90)
Week 4	113	0.58 (0.38)	0.77 (0.00-0.86)
CEQ-5D-5L			
Baseline	156	0.23 (0.21)	0.27 (0.18-0.32)
Week 4	122	0.38 (0.24)	0.30 (0.21-0.54)
DEMQOL-Proxy-U			
Baseline	225	0.63 (0.13)	0.62 (0.52-0.73)
Week 4	228	0.57 (0.26)	0.65 (0.51-0.73)

	Complete cases		
Measure	**n**	**Mean (SD)**	**Median (IQR)**
EQ-5D-5L			
Baseline	43	0.25 (0.20)	0.28 (0.19-0.34)
Week 4	43	0.58 (0.31)	0.66 (0.34-0.83)
DEMQOL-U			
Baseline	42	0.81 (0.14)	0.86 (0.70-0.91)
Week 4	42	0.80 (0.12)	0.82 (0.69-0.88)
CEQ-5D-5L			
Baseline	92	0.23 (0.22)	0.27 (0.19-0.28)
Week 4	92	0.38 (0.25)	0.29 (0.21-0.54)
DEMQOL-Proxy-U			
Baseline	184	0.64 (0.13)	0.65 (0.55-0.73)
Week 4	184	0.66 (0.12)	0.67 (0.55-0.74)

Table 3: Spearman correlation between HRQoL utilities and health status at baseline and Week 4

Baseline

Measure	n	EQ-5D-5L	n	DEMQOL-U
MMSE	82	0.22*	71	0.05
CSDD	81	-0.33**	70	-0.59**
MBI	82	0.46**	71	0.37**
PainAd	82	-0.45**	71	-0.46**
Measure	n	CEQ-5D-5L	n	DEMQOL-Proxy-U
MMSE	156	0.12	225	0.18**
CSDD	156	-0.48**	224	-0.48**
MBI	156	-0.003	225	0.21**
PainAd	156	-0.37**	225	-0.26**

Week 4

Measure	n	EQ-5D-5L	n	DEMQOL-U
MMSE	77	-0.13	81	0.22*
CSDD	77	-0.17	81	-0.34**
MBI	77	0.16	80	0.14
PainAd	77	-0.06	81	-0.002
Measure	n	CEQ-5D-5L	n	DEMQOL-Proxy-U
MMSE	122	0.15	196	0.07
CSDD	122	-0.27**	196	-0.20**
MBI	122	0.38**	196	0.21**
PainAd	121	-0.07	195	-0.15*

* correlation significant at the 0.05 level (two tailed)
** correlation significant at the 0.01 level (two tailed

Table 4a: HRQoL utilities (mean and SD) by health status at baseline

Baseline

Measure	n	EQ-5D-5L	n	DEMQOL-U
Cognitive impairment				
Mild (MMSE score >20)	17	0.32 (0.22)	12	0.82 (0.15)
Moderate (MMSE score 10-20)	45	0.20 (0.17)	49	0.78 (0.14)
Severe (MMSE score < 10)	20	0.14 (0.16)	10	0.83 (0.15)
P Value*		0.02*		0.21
Measure	n	CEQ-5D-5L	n	DEMQOL-Proxy-U
Mild (MMSE score >20)	1	0.17	17	0.66 (0.14)
Moderate (MMSE score 10-20)	36	0.27 (0.20)	77	0.64 (0.13)
Severe (MMSE score < 10)	119	0.22 (0.22)	131	0.61 (0.13)
P Value*		0.3		0.25
Measure	n	EQ-5D-5L	n	DEMQOL-U
Depression				
Non-case (CSDD score <11)	46	0.26 (0.18)	41	0.85 (0.12)
Probable (CSDD score 11-17)	22	0.18 (0.18)	22	0.73 (0.14)
Definite (CSDD score > 17)	13	0.12 (0.13)	7	0.68 (0.11)
P Value*		0.005		0.00
Measure	n	CEQ-5D-5L	n	DEMQOL-Proxy-U
Non-case (CSDD score <11)	87	0.29 (0.20)	124	0.67 (0.13)
Probable (CSDD score 11-17)	58	0.18 (0.19)	79	0.59 (0.11)
Definite (CSDD score > 17)	11	0.02 (0.25)	21	0.50 (0.10)
P Value*		0.00		0.00
Measure	n	EQ-5D-5L	n	DEMQOL-U
Self care				
Independence (MBI score 100)	0	-	0	-
Slight dependence (MBI score 91-99)	0	-	0	-
Moderate dependence (MBI score 61-90)	0	-	0	-
Severe dependence (MBI score 21-60)	5	0.48 (0.28)	7	0.90 (0.67)
Total dependence (MBI score 0-20)	77	0.19 (0.17)	64	0.78 (0.14)
P Value**		0.002		0.02
Measure	n	CEQ-5D-5L	n	DEMQOL-Proxy-U
Independence (MBI score 100)	0	-	0	-
Slight dependence (MBI score 91-99)	0	-	0	-
Moderate dependence (MBI score 61-90)	0	-	0	-
Severe dependence (MBI score 21-60)	7	0.10 (0.19)	11	0.66 (0.13)
Total dependence (MBI score 0-20)	149	0.24 (0.21)	214	0.63 (0.13)
P Value**		0.01		0.31

Measure	n	EQ-5D-5L	n	DEMQOL-U
Pain				
No Pain (PainAd score 0)	43	0.27 (0.19)	47	0.84 (0.12)
Mild (PainAd score 1-3)	29	0.16 (0.18)	21	0.71 (0.15)
Moderate (PainAd score 4-6)	9	0.10 (0.13)	3	0.70 (0.15)
Severe (PainAd score 7-10)	1	-0.02	0	-
P Value*		0.001		0.001
Measure	n	CEQ-5D-5L	n	DEMQOL-Proxy-U
No Pain (PainAd score 0)	93	0.28 (0.17)	129	0.65 (0.14)
Mild (PainAd score 1-3)	47	0.19 (0.21)	73	0.61 (0.12)
Moderate (PainAd score 4-6)	15	0.12 (0.34)	21	0.56 (0.10)
Severe (PainAd score 7-10)	1	-0.21	2	0.53 (0.03)
P Value*		0.00		0.002

Table 4b: HRQoL utilities (mean and SD) by health status at week 4

Week 4

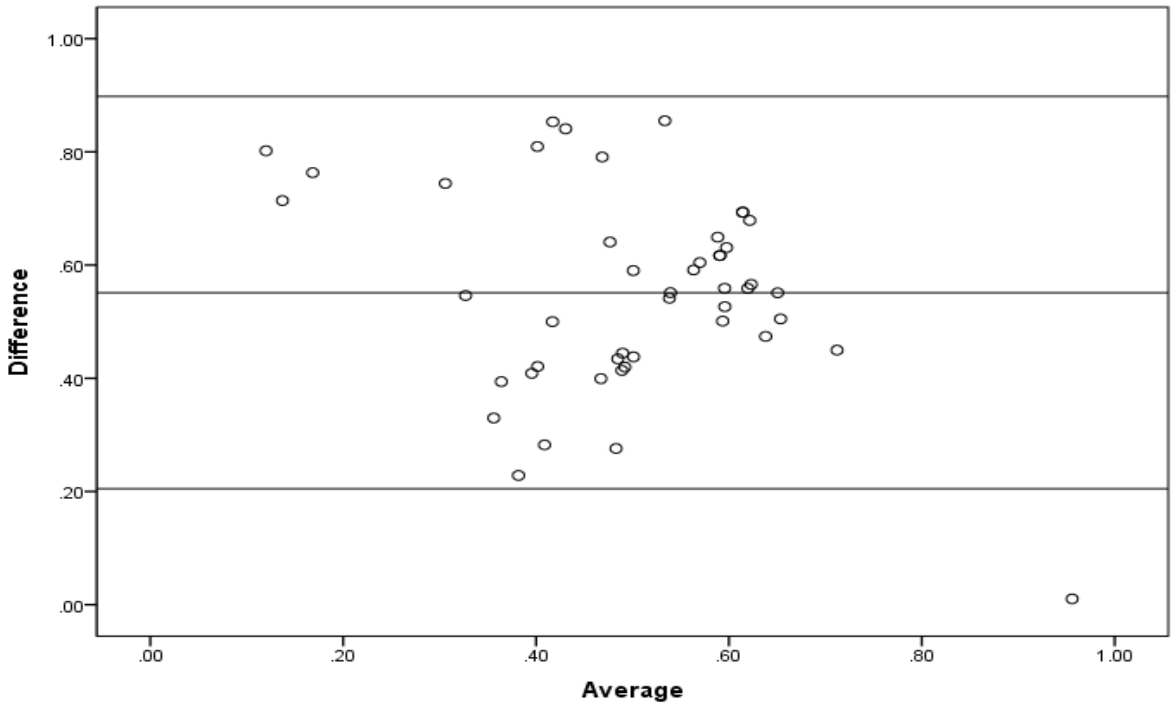
Measure	n	EQ-5D-5L	n	DEMQOL-U
Cognitive impairment				
Mild (MMSE score >20)	17	0.58 (0.26)	20	0.83 (0.10)
Moderate (MMSE score 10-20)	60	0.65 (0.30)	53	0.81 (0.11)
Severe (MMSE score < 10)	0	-	8	0.71 (0.18)
P Value*		0.27		0.22
Measure	n	CEQ-5D-5L	n	DEMQOL-Proxy-U
Mild (MMSE score >20)	3	0.44 (0.24)	19	0.71 (0.12)
Moderate (MMSE score 10-20)	15	0.49 (0.28)	74	0.66 (0.14)
Severe (MMSE score < 10)	104	0.36 (0.24)	103	0.66 (0.11)
P Value*				
Measure	n	EQ-5D-5L	n	DEMQOL-U
Depression				
Non-case (CSDD score <11)	52	0.67 (0.25)	57	0.82 (0.12)
Probable (CSDD score 11-17)	15	0.54 (0.38)	14	0.73 (0.13)
Definite (CSDD score > 17)	10	0.54 (0.34)	10	0.79 (0.11)
P Value*		0.30		0.04
Measure	n	CEQ-5D-5L	n	DEMQOL-Proxy-U
Non-case (CSDD score <11)	57	0.43 (0.25)	107	0.69 (0.12)
Probable (CSDD score 11-17)	48	0.35 (0.23)	62	0.64 (0.12)
Definite (CSDD score > 17)	17	0.28 (0.23)	27	0.62 (0.11)
P Value*		0.04		0.004
Measure	n	EQ-5D-5L	n	DEMQOL-U
Self care				
Independence (MBI score 100)	0	-	0	-
Slight dependence (MBI score 91-99)	2	0.95 (0.08)	2	0.87 (0.13)
Moderate dependence (MBI score 61-90)	19	0.67 (0.25)	19	0.81 (0.13)
Severe dependence (MBI score 21-60)	26	0.66 (0.21)	29	0.81 (0.10)
Total dependence (MBI score 0-20)	30	0.56 (0.36)	30	0.79 (0.14)
P Value*		0.28		0.90
Measure	n	CEQ-5D-5L	n	DEMQOL-Proxy-U
Independence (MBI score 100)	0	-	0	-
Slight dependence (MBI score 91-99)	0	-	2	0.84 (0.05)
Moderate dependence (MBI score 61-90)	5	0.36 (0.34)	24	0.70 (0.11)
Severe dependence (MBI score 21-60)	26	0.56 (0.25)	50	0.69 (0.13)
Total dependence (MBI score 0-20)	91	0.33 (0.21)	120	0.64 (0.11)
P Value*		0.00		0.01

Measure	n	EQ-5D-5L	n	DEMQOL-U
Pain				
No Pain (PainAd score 0)	68	0.64 (0.29)	70	0.80 (0.13)
Mild (PainAd score 1-3)	7	0.56 (0.31)	8	0.82 (0.12)
Moderate (PainAd score 4-6)	2	0.69 (0.20)	3	0.78 (0.08)
Severe (PainAd score 7-10)	0	-	0	-
P Value*		0.67		0.50
Measure	n	CEQ-5D-5L	n	DEMQOL-Proxy-U
No Pain (PainAd score 0)	96	0.38 (0.23)	162	0.67 (0.12)
Mild (PainAd score 1-3)	18	0.34 (0.27)	24	0.63 (0.11)
Moderate (PainAd score 4-6)	7	0.36 (0.29)	9	0.59 (0.13)
Severe (PainAd score 7-10)	0	-	0	-
P Value*		0.88		0.68

**Independent-samples Kruskal-Wallis Test
**Independent-samples Mann-Whitney U test

Figure 1: Bland Altman plot assessing agreement between DEMQOL-U and EQ-5D, and CEQ-5D and DEMQOL-Proxy-U utility values at baseline

(a) DEMQOL-U vs EQ-5D



(b) DEMQOL-Proxy-U vs CEQ-5D

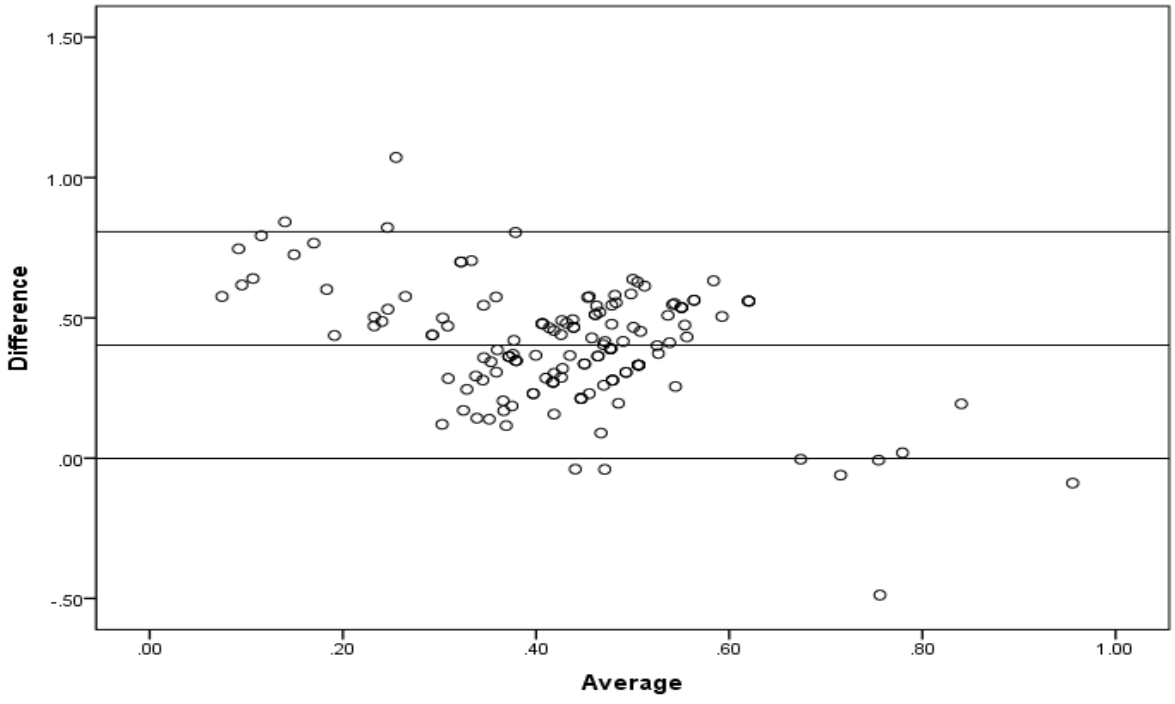
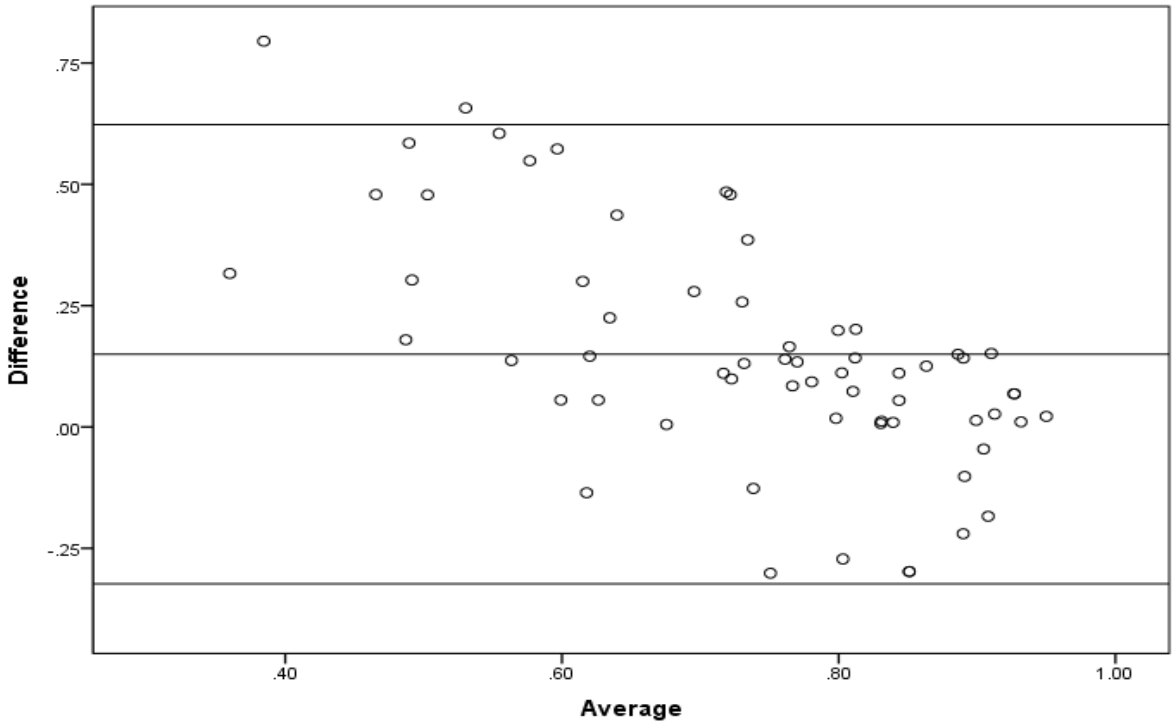


Figure 2: Agreement between DEMQOL-U and EQ-5D, and CEQ-5D and DEMQOL-Proxy-U utility values at Week 4

(a) DEMQOL-U vs EQ-5D



(b) DEMQOL-Proxy-U vs CEQ-5D