WORK IN PROGRESS PLEASE DO NOT CITE OR CIRCULATE WITHOUT THE AUTHORS PERMISSION

WHO SHOULD JUDGE QUALITY OF LIFE IN THE COGNITIVELY IMPAIRED?

AN INVESTIGATION OF PROXY COMPLETION OF EQ-5D IN PATIENTS WITH DEMENTIA

Wendy Coucill[1]
Stirling Bryan[1]
Peter Bentham[2]
Ann Buckley[2]
Ann Laight[3]

1. Health Economics Facility, University of Birmingham
2. Mental Health Services for Older Adults, Queen Elizabeth Psychiatric Hospital, Birmingham
3. The Language Unit, Park View Clinic, Birmingham

Address for correspondence:
Wendy Coucill, Health Economics Facility, University of Birmingham, Park House, 40 Edgbaston Park Road, Birmingham B15 2RT
E-mail: w.coucill@bham.ac.uk

INTRODUCTION

Dementia is recognised as a major problem in elderly populations. It is a syndrome characterised by a progressive decline in memory and cognitive functioning.[1] Alzheimer's disease is the most common cause of dementia, accounting for approximately two thirds of cases.[2,3] The prevalence of dementia is expected to rise steadily: current estimates suggest that there are 20 million sufferers world-wide, and this figure is set to double to 40 million by the year 2025 as the population ages.[4] Such predictions raise important resource allocation questions for health policy makers, not only in terms of public expenditure but also the availability and willingness of carers to provide such care. There have been recent advances, such as the development of cholinergic therapies, which have been shown in trials to provide some improvements in the cognitive function of patients.[5] These trials have, however, been viewed as limited as they have not incorporated quality of life data.[5]

Decisions concerning the funding and uptake of these new drugs can be aided by the use of economic evaluation techniques which consider alternative courses of action in terms of both their costs and benefits. One of the principal benefits to be considered in any economic evaluation of such drugs is patient health-related quality of life (HRQL).[6-8] There are two broad categories of instrument to measure HRQL: disease or function-specific and generic.[9] The latter category typically includes health profiles, such as SF-36, and preference-based measures, such as the Health Utilities Index (HUI).[9] Whilst, there have been recent developments in relation to disease-specific measures for dementia,[1, 10-16] such instruments do not enable direct comparison to be made between outcomes of various treatments for patients with different health problems.[17] This is problematic in situations where research seeks to inform resource allocation questions.[18] In contrast generic instruments have been developed for use in diverse patient populations. One group of generic instruments is that using a utility-based approach.[19]

The utility approach yields a single summary score which facilitates comparison of results across studies, and the performance of cost-utility analysis.[19] There are three utility–based measures which have been used in patients with dementia: HUI, the Quality of Well-being Scale (QWB) and the EuroQol EQ-5D instrument (EQ-5D).[20-22] The choice of instrument for use in the study reported in this paper was EQ-5D which is one of the most commonly used utility instruments.[17] This can be simply administered to patients in the form of a self-completed questionnaire and is increasingly being used in patients with neurological disorders.[23, 24] There are two core components to the instrument: a description of the respondent's 'own health' using a health state classification system with five dimensions (mobility, self-care, usual activities, pain/discomfort and anxiety/depression) and a rating of 'own health' by means of a visual analogue 'thermometer' scale (VAS).[25] Valuations

of the EQ-5D health states, have been obtained in several general population studies[26] with the most commonly applied valuation system or "tariff" being that developed in the UK.[27]

There are, however, inherent difficulties in the assessment and measurement of HRQL in patients with dementia. The major problem is that all patients with dementia have impaired cognitive abilities.[28] Due to such cognitive dysfunction, individuals with dementia may have a loss of insight and capacity for self-observation and forget recent experiences and feelings. In addition, demented patients may be unable to recognise disabilities such as impairment of memory as a direct consequence of the neurological dysfunction, referred to as anosognosia. Patients may lack insight into their illness and psychological mechanisms, such as denial, may also play an important role.[29] Therefore, in some cases patients may not be seen as reliable sources of information.[30,31] Some research suggests that patients may be able to respond to carefully designed instruments in mild stages of the disease, with instruments such as the Dementia Quality of Life scale (DQol) being specifically developed for self-administration.[32] Selai and Trimble [33] suggest that patients may also be able to discuss their quality of life in a semi-structured interview. It is, however, unclear at which point in the disease process patients are no longer able to provide accurate reports. Thus, judgements relating to HRQL are often obtained using proxies, such as carers or clinicians, which raises issues regarding the reliability and validity of such responses.

Where proxies are used, the underlying assumption made is that the proxy can report accurately on the status of the subject.[34] Whilst proxy information on past and current health problems such as diabetes, thyroid disease and stroke has been shown to be in almost complete agreement with self-report,[35] proxy ratings of HRQL do not always correlate well with the patient's own answers.[6] Zimmerman and Magaziner [34] suggest that agreement is a function of the type of information sought, the characteristics of the subject and the characteristics of proxies and their ability to observe subjects. It has also been suggested that agreement is better when questions concern attributes with observable manifestations. [36] Proxies have also been found to describe a higher level of impairment in functioning and emotional well-being, relative to subjects. This is particularly marked where the subjects have cognitive impairment.[37] Choice of proxy may also be influential upon the extent to which there is agreement. It has been suggested that health care providers (i.e. clinicians and nurses) may be better raters of patients' functional status and physical symptoms, and that 'significant others' (i.e. carers) may be more accurate in assessing patients' psychological and social health.[38] Busschbach et al[31] have suggested that head-to-head comparisons of proxies are required in order to study the level of agreement among patients and proxies.

Whilst there are many difficulties associated with the use of data from proxies, such data does allow fuller coverage to be achieved, with data then being available from subjects who are unable or unwilling to respond. This is particularly important in the evaluation of health care technologies where the exclusion of such patients could lead to bias in the estimation of the benefit associated with an intervention. For example, improvements in quality of life may be greatest for those with higher initial levels of cognitive impairment, and who are also less likely to be able to provide self-reports. Magaziner [36] urges researchers to identify the imperfections surrounding the use of proxies and to carefully document the potential error they introduce.

There are two components to the study reported in this paper. Part 1 investigates the inter-rater agreement for the EQ-5D instrument when completed by two categories of proxy (i.e. carer and clinician). This section of the study tested three key hypotheses.

1. The level of agreement will be stronger between patient and carer than between patient and clinician. The rationale for this is that carers have a closer relationship to the patient and are therefore more likely to have a stronger insight into aspects of HRQL that are important to the patient.
2. The level of agreement between patient and proxies will be stronger on the 'observable' and objective dimensions of the EQ-5D questionnaire (i.e. mobility and self–care) than 'less observable' and more subjective dimensions (i.e. anxiety/depression and pain/discomfort).
3. The level of agreement between patient and proxies will be stronger for patients with earlier dementia. The rationale for this is that these patients are more likely to provide 'valid' reports of HRQL due to being less cognitively impaired.

Part 2 of the study explores the relationship between EQ-5D scores and disease severity. Severity was assessed in terms of:

- stage of dementia severity (Clinical Dementia Rating),[39]
- cognitive function (Mini Mental State Examination),[40,41]
- depression (Geriatric Depression Scale),[42-45]
- functional ability (Bristol Activities of Daily Living Scale),[46,47]
- behavioural and psychological symptoms (Neuropsychiatric Inventory). [48]

GENERAL STUDY METHODS AND SAMPLE DETAILS

This section reports details regarding the patient sample and method of data collection which applies to both Part 1 and Part 2 of the study. The subsequent sections of the paper, reporting methods and results are reported separately for each study part.

Patient sample

The patient sample comprised individuals with a DSM-4 diagnosis of dementia of Alzheimer's type and/or vascular dementia.[49] Subjects were eligible if they had questionable, mild or moderately severe dementia, as defined by the Clinical Dementia Rating (CDR),[39] had a 'regular carer' (i.e. a carer in direct contact with the patient at least once a week), were not in residential or nursing home care, and were under the care of South Birmingham Mental Health NHS Trust. In order to avoid sample selection bias *all* potentially eligible patients were identified from hospital records. Attempts were made to contact all eligible patients/carers and where contact was made, the carer and patient were provided with information about the study and invited to participate. This research was approved by the Local Research Ethics Committee.

Data collection

Data were collected from patients, carers and 2 study clinicians (both psychiatrists). Patients and their carers were seen on two occasions. On the first visit, they were seen by a clinician in either a clinic setting or the patient's home. On the second visit they were seen by a researcher (psychologist/health service researcher) in the patient's home (with the exception of one patient who was seen in clinic).

First visit

The clinician interviewed the patient and obtained relevant clinical information in order to make a DSM-4 diagnosis of dementia, and a global severity rating of dementia using the CDR scale. The patient then completed the Mini-Mental State Examination (MMSE) and the Geriatric Depression Scale (GDS). The carer rated the patient's ability to carry out their activities of daily living using the Bristol Activities of Daily Living Scale (BADLS). In a subsample of patients (n=41), the Neuropsychiatric Inventory was also completed by the carer. Details of these instruments are outlined in Appendix A.

After this interview, the clinician completed EQ-5D as a proxy for the patient. For the purposes of this study only two components of the questionnaire were used: the health state classification system and the VAS.[25] The rationale for selecting these parts was that these represent the core elements commonly used in health service trials. This instrument was slightly amended to accommodate completion by proxies in that it read "his/her own health" rather than "your own health", and, "he/she has..." rather than "I have…", on the dimensions of the instrument. Once this procedure was completed the EQ-5D questionnaire was placed in a coded envelope and not seen by any member of the research team until after the second visit.

Second visit

A researcher visited the patient and carer and obtained background information on:

- patient age and sex,
- total number of years of education (of both patient and carer),
- relationship of carer to the patient, and
- current living arrangements (i.e. whether the patient lived alone or with a carer).

The carer completed EQ-5D on behalf of the patient using the same amended version as the clinician. Prior to completing the questionnaire the researcher instructed the carer not to show their responses to the patient or to the researcher. The carer was then asked to ensure that all questions had been completed and to place the form in a coded envelope. The researcher then gave a copy of EQ-5D to the patient and read out the instructions for administration. Where the patient was unable to self-complete, the researcher read out the questions and responses and asked the patient to give their response.

Sample characteristics

A total of 97 patients were identified as suitable for inclusion in the study. Of these, 28 patients refused to take part and 5 of the responders failed to complete all interviews as they had moved into institutional care (n=3), or refused/were not contactable (n=2). Complete data were therefore obtained on 64 patients.

Two clinicians and 2 researchers participated in this study. Of the clinician visits, 43 were carried out in a clinic setting and 21 in the patient's home. The reason for variation in setting was one of necessity in terms of maximising patient recruitment. Of the researcher visits, 63 were carried out in the patient's home and 1 in a clinic setting. The mean time between the clinician and researcher visits was 16 days.

Table 1 reports the sample characteristics and scores on the various measures of disease severity. As can be seen the majority of patients were female and lived with a carer. The majority of carers were the patient's spouse/partner. The data collected on CDR indicated that this sample covers a range of severity from 'questionable' to 'moderate' dementia. Similar findings are also apparent on the MMSE with the mean score for the whole sample falling between the identified range for mild to moderate dementia (MMSE 10-24).[22,40,41] The mean BADLS score would seem to indicate that patients within the study were experiencing impairment in performing activities of daily. However, this should be interpreted fairly cautiously as when we examine the range in scores it can be seen that some patients scored 0 (actual number scoring 0, n=3), which would indicate no impairment. The number of patients identified as cases of depressive illness varied according to the cut off point that

was employed on the GDS of either 27% (GDS>10) or 16% (GDS>13). Although the mean score of the whole sample was below the lower of these cut off points such results also should be interpreted with caution as some patients scored 0 (n= 2). Also, as outlined in Appendix A, the GDS was not designed to classify severity of symptoms although in some literature a score of 30 has been taken to indicate severe depression.[43]

The data collected from a subsample of patients (n=41) on the NPI, would seem to indicate the presence of these symptoms in patients within the study population. Again these results should also be interpreted with caution for as with the aforementioned measures some patients had scores of 0 (n=2), and there are also issues surrounding the interpretation of the overall NPI score (details outlined in Appendix A).

PART 1: METHODS AND RESULTS

Data analysis

EQ-5D data on the health state classification system were converted into health utility scores using the UK tariff.[27] Variability in utility scores between categories of respondents (i.e. patients, carers, clinicians) was explored using repeated measures analysis of variance.[50] This analysis was extended to include disease severity as a two level factor (i.e. questionable/mild or moderate dementia) enabling the influence of disease severity upon ratings of HRQL and its interaction with category of respondent to be explored.

The agreement between patient, clinician and carer HRQL scores was determined for each pair (i.e. patient-carer, patient-clinician, carer-clinician) by calculating weighted kappa scores. These scores were interpreted according to the convention outlined by Altman[51] where kappa values are associated with the following levels of agreement: less than 0.20 = poor; 0.21 to 0.4 = fair; 0.41 to 0.60 = moderate; 0.61 to 0.80 = good; 0.81 to 1.00 = very good.

These were calculated for each of the five dimensions of EQ-5D. In order to assess whether the inter-rater agreement between patients and either proxy was stronger for patients with less severe dementia the sample was divided into two groups:

- patients with questionable/ mild dementia (CDR = 0.5,1),
- patients with moderate dementia (CDR = 2).

Weighted kappa scores were calculated for each dimension within each group. However, confidence intervals around the kappa estimates were not calculated since: “In general this is not all that useful

because unless the sample is small the confidence interval will be narrow and thus not allow for much variation in interpretation."[51]

Results

EQ-5D data

Thirty-one patients (48%) classified themselves as having no problems on all 5 dimensions on the EQ-5D instrument. Of these, 1 had questionable dementia, 23 had mild dementia and 7 had moderate dementia. The clinicians classified only 2 patients as having no problems on all 5 dimensions and carers only 1 patient. Seven of the VAS scores were rated at the top of the scale, (i.e.' Best imaginable health state' = 100), all of which were rated by patients. Of these, 1 had questionable dementia, 4 had mild dementia and 2 had moderate dementia. Two of the patients rating themselves at the top of the scale also rated themselves as having no problems on all 5 dimensions of the EQ-5D instrument, these patients both had mild dementia.

Table 2 presents descriptive statistics for EQ-5D utility and VAS scores for patient-rated EQ-5D, and carer and clinician proxy-rated EQ-5D. The highest mean score was given by the patient and the lowest was given by the carer. When the sample was split according to disease severity similar results were found. Repeated measures analysis of variance revealed that utility scores varied systematically according to who provided the data (i.e. between rater variation, $F=34.19$, $p=0.0001$) and the disease severity (i.e. between levels of disease severity variation, $F=10.44$, $p=0.002$). The interaction between disease severity and category of respondent was not statistically significant ($p= 0.27$). This indicates that the impact of severity on utility scores was similar across all three categories of respondents.

Inter-rater agreement

The data reported in Table 3 show that, in comparison with patients own responses, carers reported higher levels of disability across all 5 dimensions. However, clinicians reported fewer problems on the dimensions of 'pain/discomfort' and 'anxiety/depression', compared to patients themselves. The discrepancy between patient and proxy response is particularly noticeable for the dimension 'usual activities'.

Figure 1 reports the weighted kappa scores for the whole sample. Overall, the comparison of patient responses with either proxy had kappa scores indicating no better than 'fair' agreement across all EQ-5D dimensions. The agreement between the two proxies ranged from 'poor' to 'good'. Only weak support was found for the first hypothesis: the level of agreement between patient and carer was stronger for only 2 of the 5 dimensions when compared to agreement between patient and clinician.

Similarly evidence to support the second hypothesis is weak: the level of agreement was not consistently higher for the 'observable' dimensions. Table 4 reports weighted kappa scores for the subgroup of patients with questionable/mild dementia. The comparison of patient responses with either proxy had kappa scores indicating no better than 'fair' agreement across all EQ-5D dimensions. The agreement between the two proxies ranged from 'poor' to 'moderate'. Similarly, the first and second hypotheses appear to be no more strongly supported by the data from patients with less severe disease when compared to the full data set. Table 5 reports weighted kappa scores for the subgroup of patients with moderate dementia. For these patients, the comparison of patient responses with either proxy had kappa scores indicating no better than 'fair' agreement on the four EQ-5D dimensions, the exception being 'mobility' where agreement was 'moderate'. The agreement between the two proxies ranged from 'poor' to 'very good'. The third hypothesis being tested was that the level of agreement between groups of respondents would be stronger for those patients in the early stages of dementia. The comparison of Tables 4 and 5 does not, in general, provide evidence to support this hypothesis; the only dimension for which the hypothesis is supported is 'self-care'.

It would appear from the results of Part 1 of the study that none of the three study hypotheses are strongly supported. Also it is not clear who the appropriate proxy should be when collecting HRQL data for this patient population, as different groups of proxies provide different results.

PART 2: METHODS AND RESULTS

Rationale

The focus from this point on is the comparison of EQ-5D data (i.e. health state classification system and utility scores) from both proxies with data relating to measures of disease severity. The rationale is that the analysis reported as Part 1 did not provide an indication of preferred proxy. An underlying a priori assumption being made in these comparisons is that HRQL will be lower for those patients in the more advanced stages of dementia and also for those with higher levels of dysfunction, disability and number of symptoms. In reviewing the comparisons by the carer or clinician it may be possible to identify which of these responses are 'better', in the sense that they show a stronger relationship with the clinical measures. It is expected that patients would be rated by proxies as having more problems on the appropriate dimensions of the EQ-5D health state classification system and lower overall utility scores where there is greater evidence of cognitive dysfunction, depressive symptoms, limitations in daily living and behavioural and psychological symptoms.

Data analysis

Table 6 reports the comparisons that were made between the clinical measures and proxy responses on specific dimensions of the EQ-5D. Selection of the dimensions for comparison was based on the

likelihood that these would capture the main items being measured by each instrument. For example 'mobility', 'self/care' and 'pain/discomfort' were excluded from the comparison of EQ-5D dimensions and GDS scores since the GDS does not clearly measure these dimensions.

Visual representation, namely box-and-whisker plots, was the method used to explore such relationships (see Appendix B).* The scores on each of the clinical outcome measures (*y-axis*) were plotted against the distribution of responses on the EQ-5D health state classification system *(x-axis).* The EQ-5D data were coded as follows:

- 1= no problems,
- 2=some/ moderate problems,
- 3= extreme problems or unable to perform.

The relationship between EQ-5D utility scores and the measures of disease severity were also examined statistically using Spearman's rank correlation (given the non-parametric nature of the data).

Results

Appendix B reports the results of the comparisons between scores on clinical measures and proxy responses on the specific dimensions of the EQ-5D health state classification system. The left-hand column always reports the comparison for carer-proxy data and the right-hand column for clinician-proxy data.

Scores on the MMSE scale were compared with proxy responses on the EQ-5D dimensions 'usual activities', 'self-care' and 'anxiety/depression' (see Figures B1 to B6). The boxplots reveal the expected relationship between MMSE and 'usual activities' and 'self care', with lower MMSE scores associated with higher levels of reported impairment. No such relationship is evident from the comparison of MMSE with 'anxiety/depression' scores. These findings are similar across both proxies – on the basis of these data it is not possible to identify the preferable source of proxy information.

Scores on the GDS were compared with proxy responses on the EQ-5D dimensions 'usual activities', and 'anxiety/depression' (see Figures B7 to B10). There is no clear trend apparent from the boxplots between GDS scores and carer and clinician proxy ratings on the E-5D dimension 'usual activities'. For the dimension 'anxiety/depression' higher median GDS scores on levels 2 and 3 can be seen on the clinician proxy-rated EQ-5D responses. This is less apparent on the carer proxy rated responses –

* By way of explanation of the boxplots: the box represents 50 % of the scores (25^{th}-75^{th} percentile), the central line across the box represents the median score, and the furthest point on the whiskers represent the highest and lowest values, excluding outliers and extreme values.

although the GDS scores are higher in level 3 than the preceding levels, there is little difference between the with median GDS scores for levels 1 and 2.

Scores on the BADLS scale were compared with proxy responses on the EQ-5D dimensions 'usual activities', 'self-care' and 'anxiety/depression' (see Figures B11 to B16). The boxplots indicate a positive relationship, as one would predict, between BADLS scores and EQ-5D scores for the dimensions 'self-care' and 'usual activities'. This finding clearly holds for both proxy respondents. However, this trend is not apparent on the dimension 'mobility' for either proxy, although the range on mobility scores is limited – level 3 was never indicated.

Scores on the NPI were compared with proxy responses on the EQ-5D dimensions 'usual activities', 'self-care' and 'anxiety/depression' (see Figures B17 to B22). There is substantial variation in all of these plots and clear trends are difficult to identify. Perhaps the most apparent finding emerges from the comparison of clinician proxy-rated responses on the dimension 'anxiety/depression' and NPI scores, with higher scores being seen in level 2 than level 1, as would be expected. However, this should however be interpreted with caution since there is limited variability in the EQ-5D data. The clinicians did not record any patients in level 3 (i.e. extreme anxiety/depression) and only 5 patients are in level 2 (i.e. moderate anxiety/depression).

Table 7 reports the Spearman rank correlation coefficients between the various clinical measures. Four of these measures were significantly negatively correlated with both the carer and clinician proxy rated EQ-5D utility scores. Two of these were at a significance level of $p<0.01$ (GDS and BADLS) and the remaining two at $p<0.05$ (NPI and CDR). The correlation coefficients are very similar for both proxy sources of information and so, again, a strong indication of the preferred proxy is not provided.

DISCUSSION

The study reported in this paper represents one of the first to investigate the inter-rater agreement of proxy-completion of EQ-5D in the context of patients with dementia. The patient sample upon which data were collected included patients across a broad range of disease severity, from 'questionable' to 'moderate' dementia, which allowed exploration of variation in levels of agreement by severity. This study builds upon earlier 'pilot' work reported by Selai et al.[22] They examined the level of agreement between carer (acting as proxy for the patient) and patient self-rated responses using EQ-5D and found that the level of agreement across dimensions ranged from 'fair' to 'good', with the weakest level of agreement being on the dimension 'usual activities'. The results reported here, whilst

supporting the latter point, that the weakest level of agreement was found on the dimension 'usual activities', in general indicated poorer levels of agreement ('poor' to 'fair') between carer and patient ratings of patient HRQL, than reported by Selai et al.

The principal findings of this study are that none of the three hypotheses are strongly supported. (1) The level of agreement was not found to be stronger between the patient and carer than between the patient and clinician. (2) The level of agreement between the patient and proxies was not found to be stronger on 'observable' and objective dimensions of EQ-5D, even for the sub-group of patients with questionable/mild dementia. (3) The level of agreement between patients and proxies was not found to be stronger for those with questionable/mild dementia.

The data reported in this paper relating to hypothesis 2 are not in line with the findings of Magaziner[36] who suggested that proxies are poorer judges of less observable and more private symptoms compared to symptoms with observable manifestations. The findings relating to hypothesis 3 have two possible explanations. First, given that the questionable/mild sub-sample contained only a small number of patients with 'questionable' dementia, the sub-sample may have contained patients who were unable to recognise their disability. Second, it has been suggested elsewhere that patients with mild dementia sometimes fail to fully appreciate their disability as a consequence of the carer providing help and support in a manner that serves to shield the patient.[6] This argument is supported by the observed data, which show that carers were more likely to report problems on all dimensions, compared to patients. Such a finding is consistent with the data presented in Selai et al,[22] who found that carers rated patients as having a worse quality of life than did patients themselves on all dimensions of the Quality of Life Assessment Schedule (QOLAS).

A further point of interest is that the data reported in this study indicate that the level of agreement between carers and clinicians, in acting as proxies for the patient, was particularly poor for the dimensions of 'pain' and 'anxiety'. This result could be explained by the fact that these two dimensions are not clearly observable. In addition, the agreement was only 'fair' for usual activities' which may reflect different perceptions between carers and clinicians as to what constitutes 'usual activities' for the patient in question.

Whilst the validity, sensitivity and reliability of EQ-5D has been demonstrated in a range of populations, including the elderly and patients with rheumatoid arthritis or stroke,[52-59] there is some evidence it registers a "ceiling effect" especially in general population surveys. Ceiling effects arise where best health state is still associated with substantial HRQL impairment.[60] The ceiling effect in EQ-5D is seen where many respondents classify themselves as having no problem on any of the five dimensions.[54,61] This is likely to be less of a limitation in situations where EQ-5D is completed by

patients who, by definition, have some disease burden.[61] The findings of the study reported in this paper provide some support for this general conclusion since only 18% (34/189) of 'all' (patient, carer proxy and clinician proxy-rated) observations were at the ceiling health state. Caution should, however, be exercised in the interpretation of this result since 91% (31/34) of ceiling observations were provided by patients. Indeed, 48% (31/64) of the patients in this study rated themselves as having no problems on all five dimensions of the health state classification system. Given that ceiling effects were not apparent for all raters it is unlikely that a true ceiling effect was seen. Anosagnosia is almost certainly the explanation for the high numbers of patient self-ratings classified at ceiling. If it is indeed the case that patients within this study lacked insight into their disability then the validity of these patient responses is clearly open to question.

A potential weakness in this study relates to the generalisability of the results, which is limited by the fact that only two clinicians were involved. This reflected practical limitations imposed by the resource constraints on the study.

This study raises important issues concerning the acceptability of EQ-5D in patients with dementia and the use of proxies to provide such data. On the basis of this study, a number of conclusions can be drawn. Firstly, we would suggest that our data provide some support for the use of EQ-5D in this patient population when interviewer administered. The basis for this assertion is that the number of missing responses on the EQ-5D questionnaire was very small; only 3 of the carer proxy-ratings on the health state classification system were missing and only a single clinician proxy rating on the VAS was missing.

Secondly, in relation to who is likely to be the most useful source of information, we would suggest that our data raise serious concerns about the validity of patient self-rating, even for those with mild dementia. The rationale for this is that patient self-rated responses accounted for 91% (31/34) of all ceiling (i.e. 'full health') responses. As indicated earlier, ceiling effects were not apparent for all raters and so it is unlikely that the data reflect a true ceiling effect was seen. However, researchers should give careful consideration to the choice of proxy for the collection of HRQL data from patients with dementia, as the use of different groups of proxies can provide different results.

Thirdly, despite the further analysis that was conducted in Part 2 of the study, we were not able to identify the preferred proxy source of information on health status for patients with dementia. Although it was apparent for some of the dimensions that there was clearly a relationship between proxy responses on the EQ-5D health classification system, this did not allow us to identify in any consistent manner whether carers or clinicians represented better raters. What was apparent, was that carers tended to report higher levels of disability than clinicians, this is a trend that has been reported

in previous literature.[37] We would suggest, therefore, that researchers collect data from both carers and clinicians acting as proxies, rather than relying upon a single proxy to collect such data. It is unclear from our data who the appropriate proxy should be, and whether or not carers represent better proxies than clinicians.

Finally an inverse relationship between EQ-5D utility scores and some measures of disease severity was identified, a pattern that is seen for both clinician and carer proxies. This would seem to suggest that at a group level both proxies rate patients as having poorer overall HRQL when: (1) stage of dementia severity increases; (2) there are greater limitations in activities of daily living (3) there is a higher presence of depressive symptoms (4) there is a higher presence of behavioural and psychological symptoms. Of particular interest is that the MMSE scores were not correlated with either the carer or clinician proxy reported EQ-5D utility scores. There are two potential explanations for this for this occurrence. The first being that the MMSE is a relatively simple and brief instrument, so that the level of executive function that this measures may not impact significantly upon older patients HRQL. This is a finding that was also shown by Logsdon with the QOL-AD scale.[62] The second explanation is that the EQ-5D instrument does not possess a cognitive dimension and that the impact of cognitive decline on HRQL is not clearly captured within the other dimensions of the instrument. It would, however, be difficult to envisage a scenario whereby cognitive dysfunction has no impact on any other aspect of HRQL.

In conclusion, a number of questions remain unanswered and require further research. For example, it is not currently clear who is the appropriate proxy (carer or health professional) when collecting HRQL data using EQ-5D in a population of patients with dementia. In addition, uncertainty remains in terms of the levels of dementia severity at which patients are able to provide 'valid' ratings of health-related quality of life using the EQ-5D instrument. Finally, the generalisability of the results reported here should be explored, in terms of a broader age range and severity range of patients with dementia and across a larger cohort of clinicians.

References

1. Walker MD, Salek SS, Bayer AJ. A review of quality of life in Alzheimer's disease. Part 1: issues in assessing disease impact. Pharmacoeconomics 1998;14 (5):499-530.

2. Stewart AL, Sherbourne CD, Brod M. Measuring health-related quality of life in older and demented populations. In: Spilker B, editor. Quality of Life and Pharmacoeconomics in Clinical Trials, 2nd edition, Philadelphia: Lippincott-Raven Publishers; 1996:819-829.

3. Bosanquet N, May J, Johnson N. Alzheimer's disease in the United Kingdom, burden of disease and future care. Health Policy Review paper no 12;1998.

4. The Wellcome Trust Research directions in Alzheimer's disease. Wellcome News Supplement Q3(S),1998.

5. Winblad B, Wimo A. Assessing the societal impact of acetylcholinesterase inhibitor therapies. Alzheimer Disease and Associated Disorders 1999;13 (Suppl. 2):9-19.

6. Selai C and Trimble MR. Assessing quality of life in dementia. Aging and Mental Health 1999;3 (2):101-111.

7. Kelly CA, Harvey RJ, Cayton H. Drug treatments for Alzheimer's disease. BMJ 1997; 314:693-694.

8. Burns A. Alzheimer's disease; pharmacological developments to the year 2000. Human Psychopharmacology 1995;13:21-34.

9. Guyatt GH, Feeny DH, Patrick DL. Measuring health-related quality of life. Ann Intern Med 1993;1188:622-629.

10. Brod M, Stewart AL, Sands L. Conceptualisation of quality of life in dementia. Journal of Mental Health and Aging 1999;5(1):7-19.

11. Albert SM, Castillo-Castanada C, Jacobs DM, Sano M, Bell K, Merchant C. et al. Proxy-reported quality of life in Alzheimer's patients: comparison of clinical and population based samples. Journal of Mental Health and Aging 1999;5(1):49-58.

12. Rabins PV, Kasper JD, Kleinmann L, Black BS, Patrick DL. Concepts and methods in the development of the ADRQL: an instrument for assessing health-related quality of life in person's with Alzheimer's disease. Journal of Mental Health and Aging 1999;5(1):33-48.

13. Logsdon RG, Gibbons LE, McCurry SM, Teri L. Quality of life in Alzheimer's disease: patient and caregiver reports. Journal of Mental Health and Aging 1999;5(1):21-32.

14. Salek MS, Schwartzberg E, Bayer AJ. Evaluating health-related quality of life in patients with dementia: development of a proxy-administered questionnaire [abstract]. Pharm World Sci 1996;18 (5 Suppl. A):6.

15. Salek SS, Walker MD, Bayer AJ. The community dementia quality of life profile (CDQLP): A factor analysis. Quality of Life Research 1999;8(7):660.

16. Dejong R, Osterlund OW, Roy MS. Measurement of quality-of-life changes in patients with Alzheimer's disease. Clinical Therapeutics 1989;11(4):545-554.

17. Fitzpatrick R, Davey C, Buxton MJ, Jones DR. Evaluating patient-based outcome measures for use in clinical trials. Health Technology Assessment (Review) 1998;2(14).

18. Cairns J. Measuring health outcomes (editorial). BMJ 1996;313:6.

19. de Haan, MS, Aaronson N, Limburg M, Langton-Hewer R, van Crevel H. Measuring quality of life in stroke. Stroke 1993;24(2):320-327.

20. Kerner D, Paterson TL, Grant I, Kaplan RM. Validity of the Quality of Well-Being Scale for patients with Alzheimer's disease. Journal of Aging and Health 1998;10(1):44-61.

21. Neumann PJ, Kuntz KM, Leon JL, Araki SS, Hermann RC, Hsu M, Weinstein MC. Health Utilities in Alzheimer's disease: a cross sectional study of patients and caregivers. Med Care 1999;37(1):27-32.

22. Selai CE, Trimble M, Rossor M, Harvey RJ. The Quality of Life Assessment Schedule (QOLAS): a new method for assessing quality of life (QOL) in dementia; In: Logsdon R, Albert S, editors. Assessing quality of life in dementia. New York: Springer; 2000:31-48.

23. Dorman PJ, Wadell F, Slattery J, Dennis M, Sandercock P. Is the EuroQol a valid measure of health-related quality of life after stroke? Stroke 1997; 28(10):1876-1882.

24. Gorter JW. Anticoagulants in the secondary prevention of stroke. 1999. Thesis University of Utrecht, ISDN: 90-393-2231-7.

25. Brooks R, with the EuroQol Group. EuroQol: the current state of play. Health Policy 1996;37:53-72.

26. Coons SJ, Rao S, Keinnninger DL, Hays RD. A comparative review of generic quality-of-life instruments. Pharmacoeconomics 2000;17(1):13-35.

27. Dolan P. Modelling valuations for EuroQol health states. Med Care 1997;35(11):1095-1108.

28. Whitehouse PJ. Pharmacoeconomics of dementia. Alzheimer Disease and Associated Disorders 1997;11(Suppl.5):S22-S33.

29. Whitehouse PJ, Orgogozo JM, Becker RE, Gauthier S, Pontecorvo M, Erzigkeit H et al. Quality of life assessment in dementia drug development. Position Paper from the International Working Group on Harmonization of Dementia Drug Guidelines. Alzheimer Disease and Associated Disorders 1997;11(Suppl. 3): S56-S60.

30. Logsdon RG, Albert SM. Assessing quality of life in Alzheimer's disease: conceptual and methodological issues. Journal of Mental Health and Aging 1999;5(1):3-6.

31. Busschbach JJV, Brouwer WBF, van der Donk A, Passchier J, Rutten FFH. An outline for a cost effectiveness analysis of a drug for patients with Alzheimer's disease. Pharmaceconomics 1998;13 (1 pt 1):21-34.

32. Brod M, Stewart AL, Sands L, Walton P. Conceptualisation and measurement of quality of life in dementia: the dementia quality of life instrument (DQoL). Gerontologist 1999 Feb; 39 (1): 25-35.

33. Selai C. and Trimble M. Quality of life (QOL) in dementia. Quality of Life Newsletter 19:1. Mapi Research Institute. Lyon, 1998.

34. Zimmerman SI, and Magaziner J. Methodological issues in measuring the functional status of cognitively impaired nursing home residents: the use of proxies and performance based measures. Alzheimer Disease and Associated Disorders 1994;8 (Suppl.1): S281-S290.

35. The Medical Research Council Cognitive Function and Ageing Study. Survey into health problems of elderly people: a comparison of self-report with proxy information. International Journal of Epidemiology 2000;29:684-697.

36. Magaziner J. Use of proxies to measure health and functional outcomes in effectiveness research in persons with Alzheimer disease and related disorders. Alzheimer Disease and Associated Disorders 1997;(Suppl.6):168-174.

37. PJ Neumann, SS Araki, M Gutterman. The use of proxy respondents in studies of older adults: lessons, challenges and opportunities. Journal American Geriatrics Society 2000;48:1646-1654.

38. Sprangers MAG. and Aaronson NK. The role of health care providers and significant others in evaluating the quality of life of patients with chronic disease: a review. Journal of Clinical Epidemiology 1992;45:743-760.

39. Hughes CP, Berg L, Danziger WL, Coben LA, Martin RL. A new clinical scale for the staging of dementia. British Journal of Psychiatry 1982;140:566-572.

40. Folstein MF, Folstein SE, McHugh PR. Mini-mental state: a practical method of grading the cognitive state of the patient for the clinician. Journal of Psychiatric Research 1975; 12: 189-98.

41. Lezak, M.D. Neuropsychological Assessment. Third Edition. Oxford University Press;1995.

42. Yesavage J, and Brink TL. Development and validation of a geriatric depression screening scale: a preliminary report. Journal of Psychiatric Research 1983; 17: 37-49.

43. Gilley DW, and Wilson RS. Criterion-related validity of the Geriatric Depression Scale in Alzheimer's disease. Journal of Clinical and Experimental Neuropsychology 1997;19(4): 489-499.

44. Burke WJ, Houston MJ, Boust SJ, Roccaforte WH. Use of the Geriatric Depression Scale in Dementia of the Alzheimer Type. Journal of the American Geriatrics Society 1989;37: 856-860.

45. Montorio I, and Izal M. The Geriatric Depression Scale: A review of its development and utility. International Psychogeriatrics 1996;8 (1):103-112.

46. Bucks RS, Ashworth D, Wilcock GK, Siegfried K. Assessment of daily living in dementia: development of the Bristol Activities of Daily Living Scale. Age and Ageing 1996; 25:113-20.

47. Byrne LMT, Wilson PMA, Bucks RS, Hughes AO, Wilcock GK. The sensitivity to change over time of the Bristol Activities of Daily Living Scale in Alzheimer's disease. International Journal of Geriatric Psychiatry 2000; 15:656-661.

48. Cummings JL, Mega M, Gray K, Rosenberg-Thompson S, Carusi D, Gornbein J. The Neuropsychiatric Inventory: Comprehensive assessment of psychopathology in dementia. Neurology 1994; 44:2308-2314.

49. American Psychiatric Association Diagnostic and statistical manual of mental disorders 4th ed. (DSM-4). Washington, DC: APA; 1994.

50. Armitage P, Berry G. Statistical methods in medical research. 2nd ed. Oxford; Blackwell Scientific Publications;1987.

51. Altman DG. Practical statistics for medical research. London: Chapman and Hall; 1991.

52. Agt HME v, Essinck-Bot M-L, Krabbe PFM, Bonsel GJ. Test-retest reliability of health state valuations collected with the EuroQol questionnaire. Soc Sci Med 1994;39:1537-1544.

53. Dorman P, Slattery J, Farell B, Dennis M, Sandercock P. Qualiatative comparison of the reliability of health state assessments with the EuroQol and SF-36 questionnaires after stroke. United Kingdom Collaborators in the International Stroke Trial. Stroke 1998;29(1):63-68.

54. Hurst NP Kind P, Ruta D, Stubbings A. Measuring health-related quality of life in rheumatoid arthritis; validity, responsiveness and reliability of EuroQol (EQ-5D). British Journal of Rheumatology 1997;36:551-559.

55. Brazier J, Jones N, Kind P, Testing the validity of the EuroQol and comparing it with the SF-36 health survey questionniare. Quality of Life Research 1993;2:169-190.

56. Essinck-Bot M-L, Krabbe PFM, Gouke JB, Aaronson NK. An empirical comparison of four generic health status measures: The Nottingham Health Profile, the Medical Outcomes Study SF-36- Item short form health survey, the COOP/WONCA charts and the EuroQol instrument. Med Care 1997;35:522-537.

57. Hurst NP, Jobanputra P, Hunter M., Lambert M, Lochead A, Brown H. Validity of EuroQol: a generic health status instrument in patients with rheumatoid arthritis. British Journal Rheumatology 1994;33:655-62.

58. BrazierJ, Walters SJ, Nicholl JP, Kohler B. Using the SF-36 and EuroQol on an elderly population. Quality of Life Research 1996;5;195-204.

59. Coast J, Peters TJ, Richards SH, Gunnell DJ. Use of the EuroQol among elderly acute care patients. Quality of Life Research 1998;7(1):1-10.

60. Gordon H, Guyatt RJ, Feeny DH, Patrick DL. Measurements in clinical trials: choosing the right approach. In: Spilker B, editor. Quality of Life and Pharmacoeconomics in Clinical Trials, 2nd edition, Philadelphia: Lippincott-Raven Publishers; 1996, p. 42.

61. Johnson. JA, Coons SJ. Comparison of the EQ-5D and SF-12 in an adult U.S. sample. Quality of Life Research 1998; 7:155-166.

62. Logsdon RG, Gibbons LE, McCurry SM, Teri L. Quality of life in Alzheimer's disease: Patient and Caregiver Reports. Journal of Mental Health and Ageing 1999;5(1):21-32.

63. Coucill W, Bryan S, Bentham P, Buckley A, Laight A. EQ-5D in dementia: an investigation of inter-rater agreement. Med Care 2001; 39(8):760-771.

Table 1: Sample characteristics

No of patients	64
Males **Females**	28 36
Age (all patients) Mean Range	 76 53-91
Clinical Dementia Rating 0.5 (questionable dementia) 1.0 (mild dementia) 2.0 (moderate dementia)	 3 36 25
Mini Mental State Examination Mean score (SD) Range	 18 (5.8) 3-26
Bristol Activities of Daily Living Scale Mean score Range	 15.8 (10.3) 0-37
Geriatric Depression Scale No of cases if GDS>10 No of cases if GDS>13 Mean score (SD) Range	 17 13 7.7 (5.3) 0-25
Neuropsychiatric Inventory (data only available on a subsample n=41) Mean score (SD) Range	 18.8 (12.7) 0-50
Living arrangements Lived with carer (who participated in study) Lived alone Lived with carer (other than carer which participated in study)	 45 16 3
Relationship to carer Spouse Adult children Siblings or relation by marriage	 42 14 8

Table 2: EQ-5D utility and 'thermometer' scores

EQ-5D Scores	*All Respondents (n=64)*			*Questionable/ mild dementia (n=39)*			*Moderate dementia (n=25)*		
	Patient self-rated	*Clinician proxy- rated*	*Carer proxy-rated*	*Patient self-rated*	*Clinician proxy- rated*	*Carer proxy-rated*	*Patient self-rated*	*Clinician proxy-rated*	*Carer proxy-rated*
Utility scores	*n*=64	*n*=64	*n*=60*	*n*=39	*n*=39	*n*=36 *	*n=25*	*n*=25	*n*=24*
Mean (SD) Range	0.80 (0.28) -0.18-1.00	0.76 (0.20) 0.19-1.00	0.60 (0.29) -0.09-1.00	0.86 (0.23) 0.09-1.00	0.81 (0.14) 0.26-1.00	0.69 (0.19) -0.07-1.00	0.72 (0.32) -0.18-1.00	0.67 (0.25) 0.19-0.88	0.46 (0.35) -0.09-0.88
Thermometer scores	*n*= 64	*n*=63*	*n*=64	*n*=39	*n*=38 *	*n*=39	*n*=25	*n*=25	*n*=25
Mean (SD) Range	71.5 (20.4) 28-100	66.6 (12.7) 30-85	60.6 (18.3) 20-95	74.7 (18.3) 40-100	72.1 (9.8) 50-85	64.0 (17.4) 30-95	66.4 (22.8) 28-100	58.2 (12.3) 30-75	55.5 (18.9) 20-90

*n<64 due to missing data n<39 due to missing data *n<25 due to missing data

***Table 3: Frequency and percentage of responses across each EQ-5D dimension**

	Patient	**Carer-proxy**	**Clinician-proxy**
Mobility			
No problems (1)	50 (78%)	39 (63%)	49 (77%)
Problems (2)	13 (20%)	23 (37%)	15 (23%)
Confined to bed (3)	1 (2%)	0 (0%)	0 (0%)
Self Care			
No problems (1)	52 (81%)	38 (59%)	45 (70%)
Some problems (2)	11 (17%)	24 (38%)	19 (30%)
Unable to wash/dress self (3)	1 (2%)	2 (3%)	0 (0%)
Usual Activities			
No problems (1)	44 (69%)	12 (19%)	3 (5%)
Some problems (2)	18 (28%)	41 (64%)	53 (83%)
Unable to perform (3)	2 (3%)	11 (17%)	8 (12%)
Pain/discomfort			
No pain/discomfort (1)	42 (66%)	30 (47.6%)	54 (84%)
Moderate pain/discomfort (2)	20 (31%)	30 (47.6%)	10 (16%)
Extreme pain/discomfort (3)	2 (3%)	3 (4.8%)	0 (0%)
Anxiety/Depression			
Not anxious/depressed (1)	43 (67%)	20 (32%)	51 (79%)
Moderately anxious/depressed (2)	19 (30%)	34 (54%)	12 (19%)
Extremely anxious/depressed (3)	2 (3%)	9 (14%)	1 (2%)

Table 4: Weighted kappa and strength of agreement for patients with questionable and mild dementia (n=39)

EQ-5D Dimension	*Patient and Carer*		*Patient and Clinician*		*Carer and Clinician*	
	Kappa	*Strength*	*Kappa*	*Strength*	*Kappa*	*Strength*
Mobility	0.13	Poor	0.35	Fair	0.47	Moderate
Self-Care	0.32	Fair	0.26	Fair	0.40	Fair
Usual Activities	0.004	Poor	-0.02	Poor	0.28	Fair
Pain	0.34	Fair	0.20	Poor	0.23	Fair
Anxiety	0.17	Poor	0.27	Fair	0.002	Poor

Table 5: Weighted kappa and strength of agreement for patients with moderate dementia (n= 25)

EQ-5D Dimension	*Patient and Carer*		*Patient and Clinician*		*Carer and Clinician*	
	Kappa	*Strength*	*Kappa*	*Strength*	*Kappa*	*Strength*
Mobility	0.43	Moderate	0.41	Moderate	0.82	Very Good
Self-Care	0.16	Poor	0.20	Poor	0.45	Moderate
Usual Activities	0.11	Poor	0.05	Poor	0.29	Fair
Pain	0.35	Fair	0.17	Poor	0.05	Poor
Anxiety	0.22	Fair	0.39	Fair	0.26	Fair

Table 6: Comparisons of clinical measures and EQ-5D dimensions

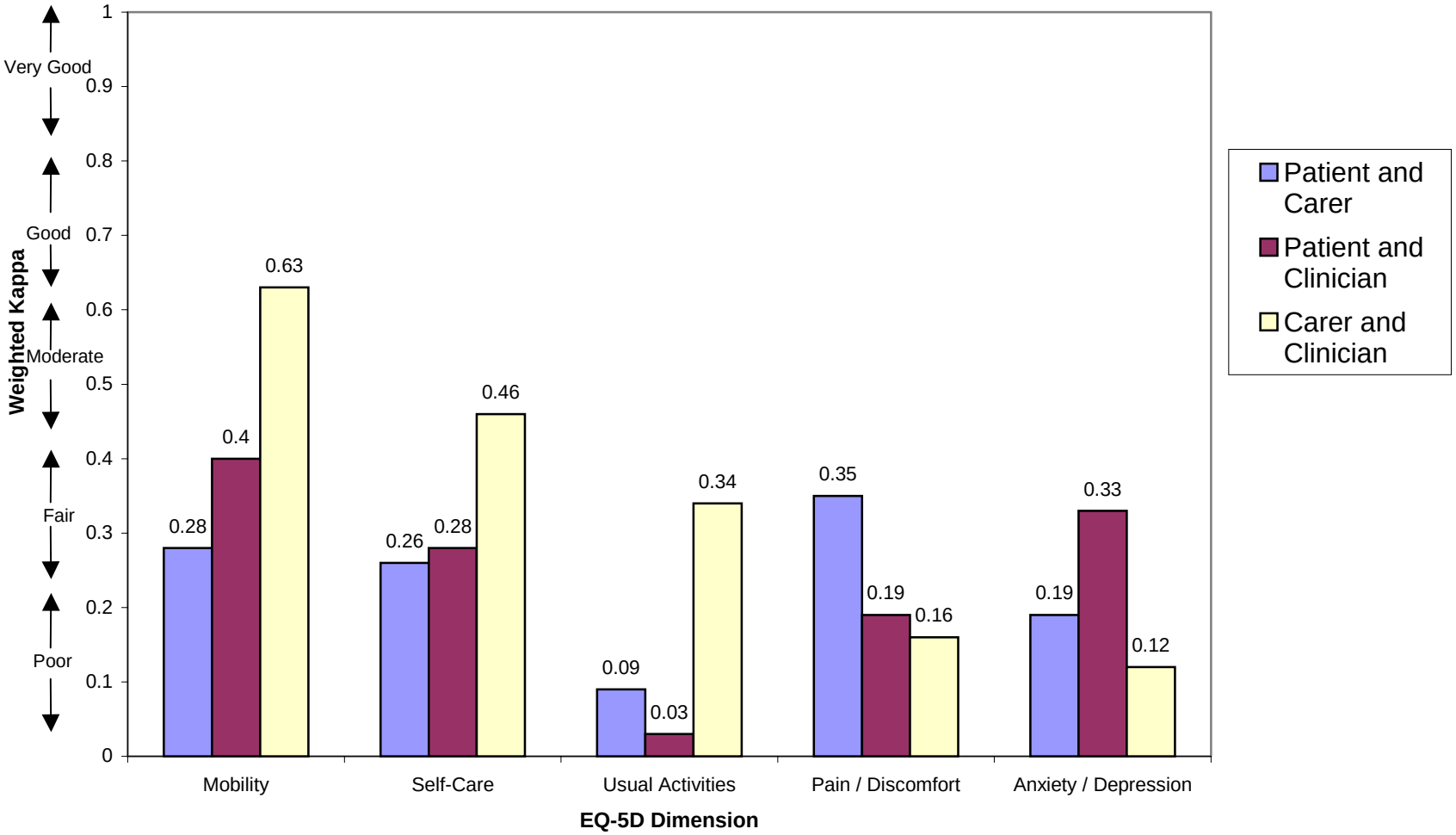
	EQ-5D Dimensions				
	Mobility	*Self-Care*	*Usual Activities*	*Pain/ Discomfort*	*Anxiety/Depression*
MMSE		*	*		*
GDS			*		*
BADLS	*	*	*		*
NPI		*	*		*

Table 7: Correlations between EQ-5D utility scores and clinical measures

	Carer proxy rated utility scores	**Clinician proxy rated utility scores**
	Spearman's correlation coefficient (r)	Spearman's correlation coefficient (r)
MMSE	0.227	0.123
GDS	-0.464**	-0.385**
BADLS	-0.679**	-0.670**
NPI	-0.401*	-0.344*
CDR	-0.264*	-0.297*

* $p<0.05$
** $P<0.01$

Figure 1. Inter-rater agreement: Weighted kappa and level of agreement for all of sample (n=64)



APPENDIX A: Details of Clinical Measures

Clinical Dementia Rating Scale (CDR)

This scale measures stage of dementia severity. It is clinician rated and interview based. Subjects are assigned a rating of healthy (CDR=0), mild (CDR=1), moderate (CDR=2), or severe dementia (CDR=3). A rating of questionable dementia (CDR=0.5) is included for patients who are neither clearly demented nor healthy. The instrument has six domains: memory; orientation; judgement and problem solving; community affairs, home and hobbies and personal care. The overall CDR is derived from scores on each of these domains, which are combined according to established rules. For details of this scoring system we would refer the reader to Hughes et al.[39]

Mini-Mental State Examination (MMSE)

This is a widely used brief screening instrument for dementia. This is a well established measure of cognitive function in elderly people. The score goes from 0-30,with the lower end of the scale representing greater cognitive impairment. Scores below 24 are considered abnormal and this is the cut off used for dementia although higher cut-off scores have been recommended for some conditions and a cut-off score of 25 has been recommended for well-educated Alzheimer patients. [22,40,41]

Geriatric Depression Scale

This scale was developed as a screening test for depression in the elderly and is designed to distinguish elderly depressives from normals. It is patient rated and interview based. It consists of 30 questions which to which the respondent answers yes or no. There are 20 positively coded items and 10 negatively coded. There is some debate about the cut off score that is used to identify case/non case. This ranges from a score of 11 or more to a more restrictive cut off of 14.[42-45] It was not designed to classify severity of symptoms although in some literature a score of 30 has been taken to indicate severe depression.[43]

Bristol Activities of Daily Living Scale (BADLS)

This scale was developed specifically for use with people with dementia It is a carer- rated instrument, which consists of 20 daily living abilities. It was designed to provide a baseline assessment of ability of demented subjects, be sensitive to change and to be brief. [46] Using this instrument patients can be assigned a minimum possible score of 0 (totally independent), and a maximum score of 60 (totally dependent). It has been suggested that this instrument is

sensitive to change, is capable of detecting minimal to severe levels of dependency in activities of daily living ability and to correlate well with different levels of cognitive performance. [46,47]

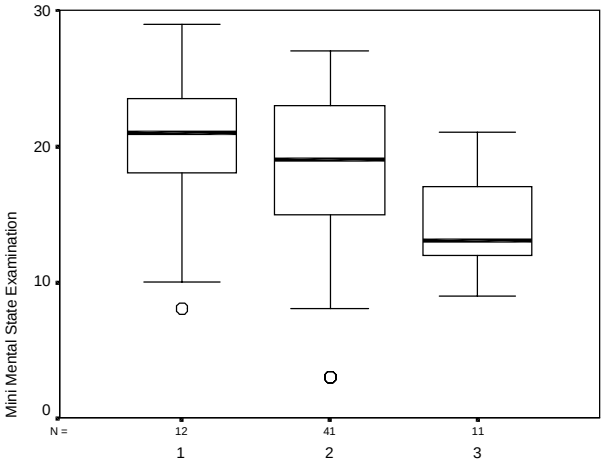
Neuropsychiatric Inventory

This is a modified version of the original instrument which assesses 12 behavioural disturbances occurring in dementia patients: delusions, hallucinations, dysphoria, anxiety, agitation/aggression, euphoria, disinhibition, irritability/lability, apathy and aberrant motor behaviour, sleep and appetite and eating disorders. It is carer rated and interview based. The NPI uses a screening strategy and scores only those domains with positive responses to screening questions. If the carer indicates that an abnormal behaviour is present they are then asked to rate the severity (1=mild,2=moderate,3=severe) and frequency (1=occasionally,less than once per week, 2=often, about once a week, 3=frequently,several times per week but less than every day; 4= very frequently, once or more per day or continuously) of the behaviour.[48] Scores are calculated by multiplying subscale scores by frequency and severity. A global score can be generated by summing the total scores, thus giving a maximum score of 144. Cummings et al [48] suggest that as the global score sums many disparate behaviours it should be regarded as an imprecise guide to the total behavioural disturbance exhibited by the patient**.**

Appendix B: Box-and-whisker plots comparing EQ-5D and disease severity data

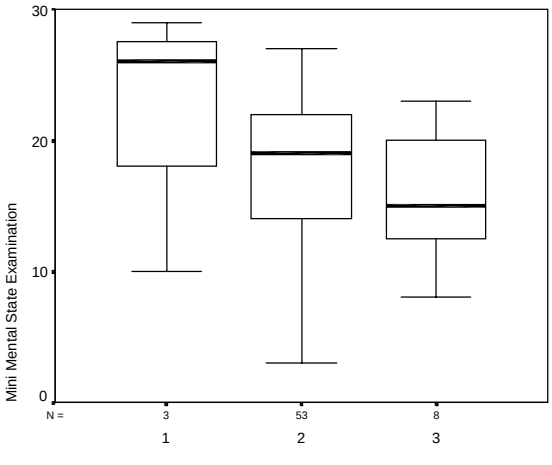
MMSE scores and EQ-5D dimension comparisons

B1



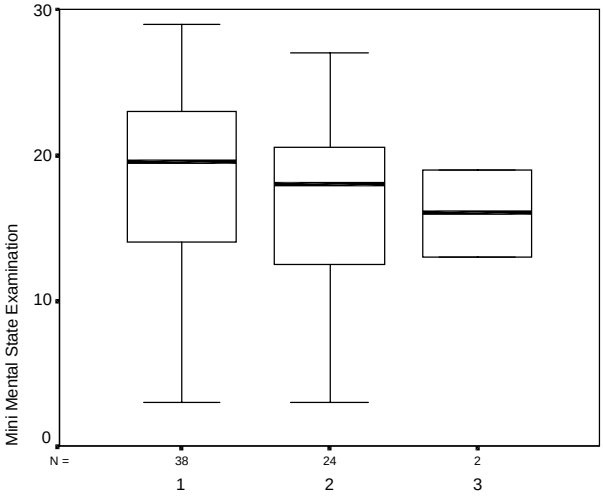
Carer proxy-rated EQ-5D 'usual activities'

B2



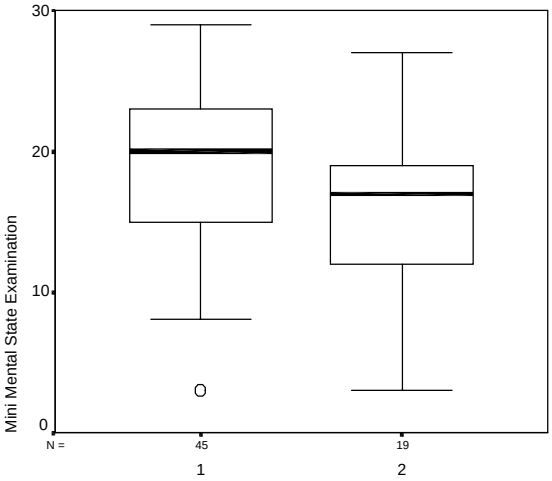
Clinician proxy-rated EQ-5D'usual activities'

B3



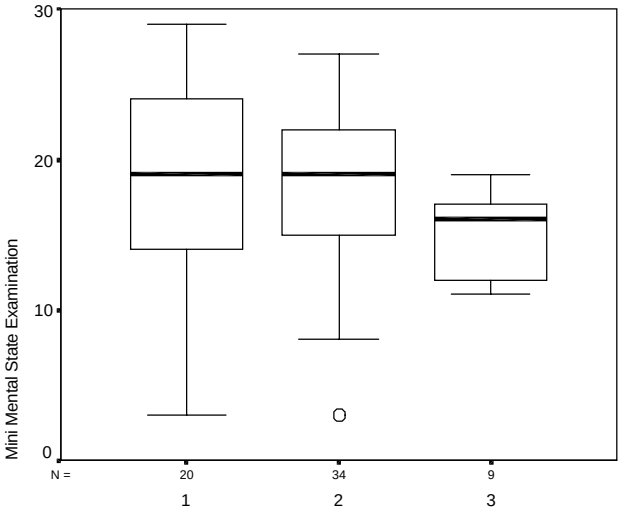
Carer proxy-rated EQ-5D 'self-care'

B4



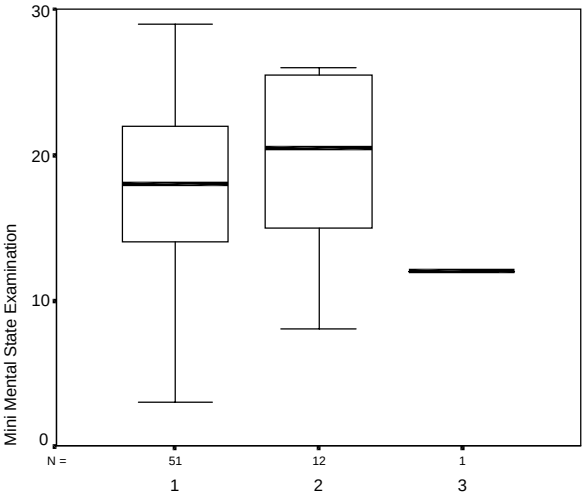
Clinician proxy-rated EQ-5D 'self-care'

B5



Carer proxy-rated EQ-5D 'anxiety/depression'

B6

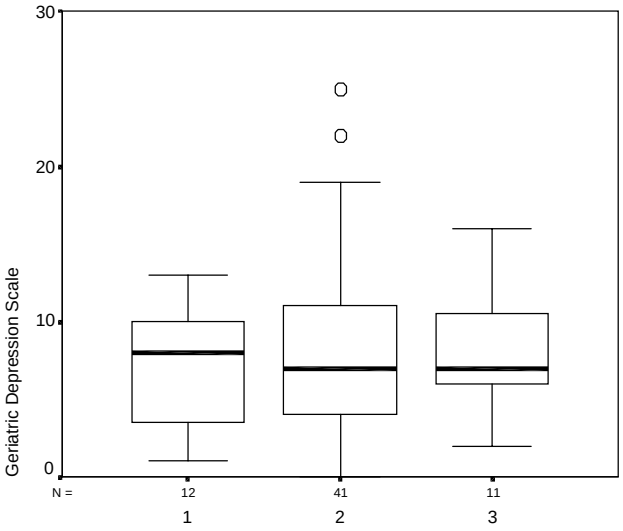


Clinician proxy-rated EQ-5D 'anxiety/depression'

o = outliers
*=extreme values

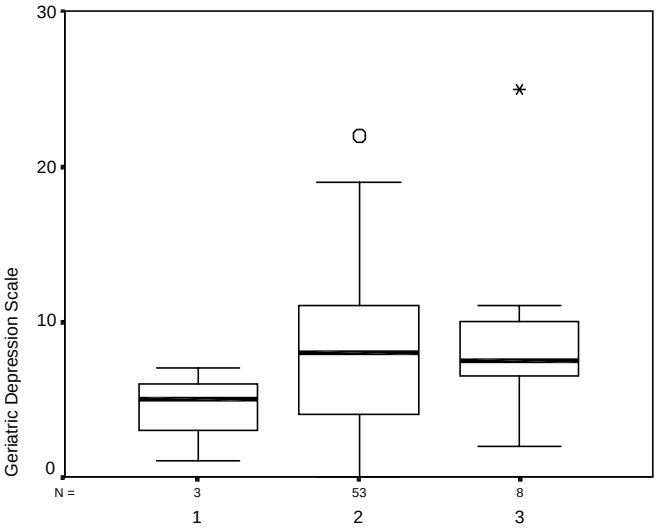
GDS scores and EQ-5D dimension comparisons

B7



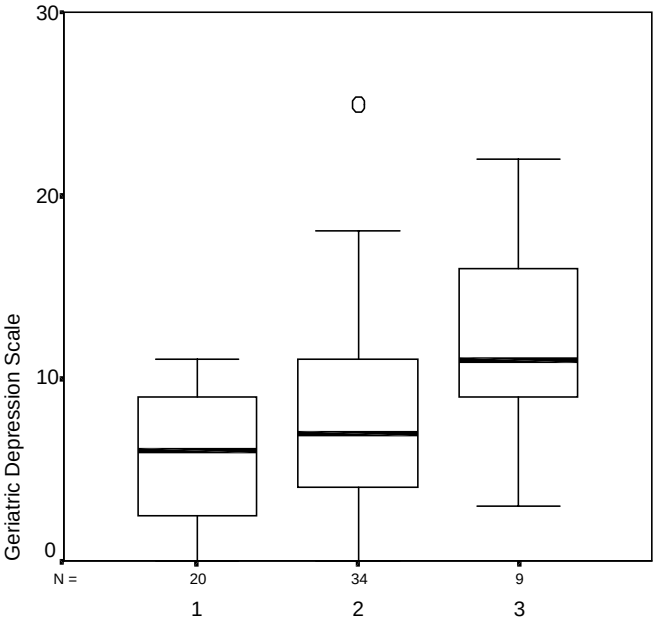
Carer proxy-rated EQ-5D 'usual activities'

B8



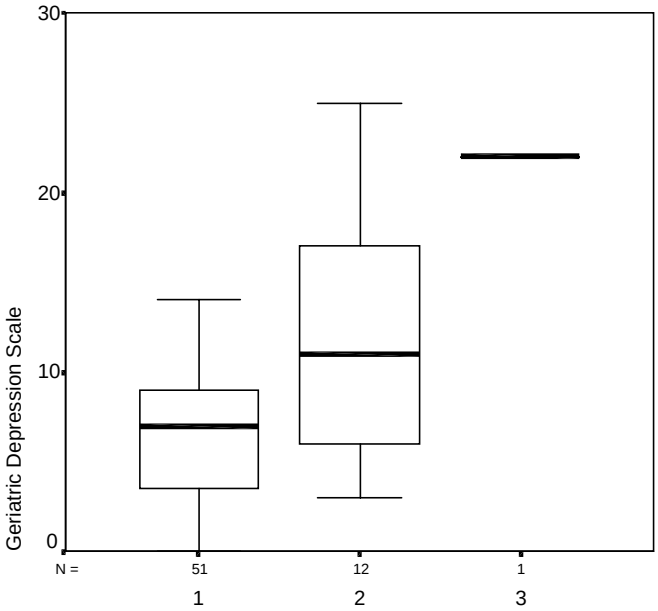
Clinician proxy-rated EQ-5D 'usual activities'

B9



Carer proxy-rated EQ-5D 'anxiety/depression'

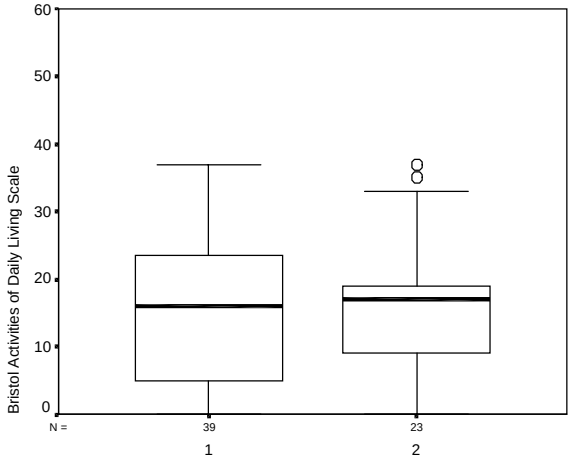
B10



Clinician proxy-rated EQ-5D 'anxiety/depression'

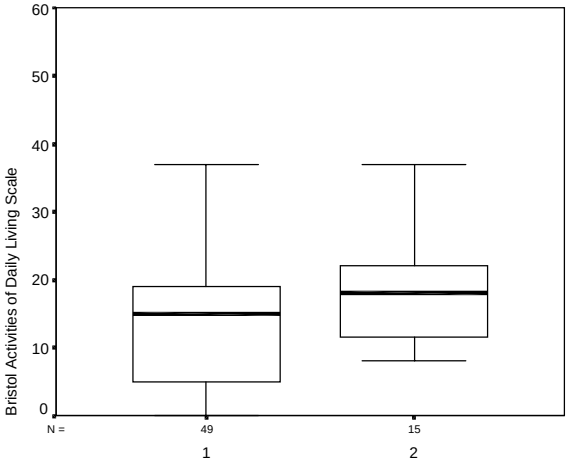
BADLS and EQ-5D dimension comparisons

B11



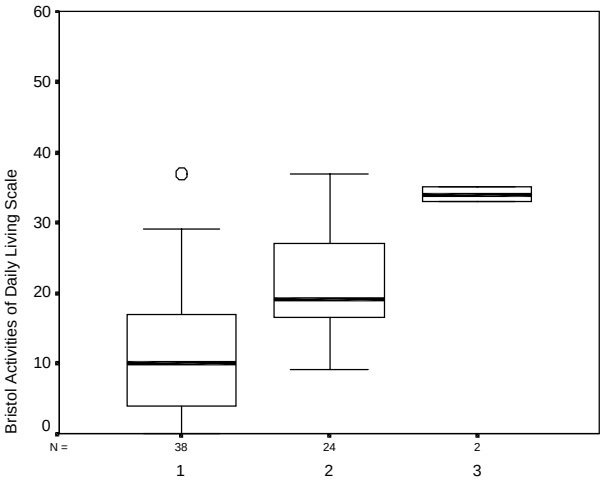
Carer proxy-rated EQ-5D 'mobility'

B12



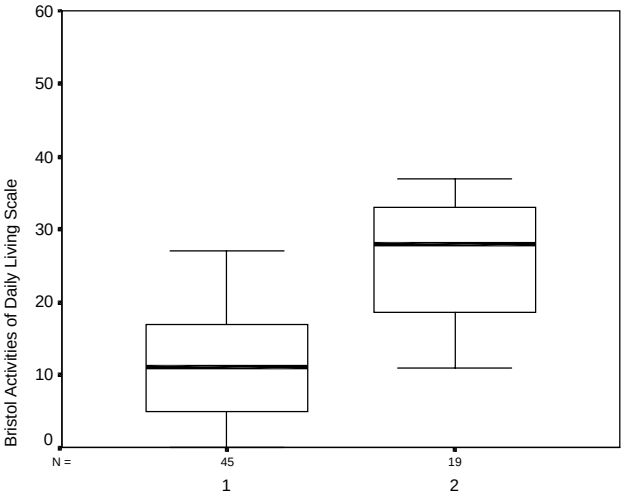
Clinician proxy-rated EQ-5D 'mobility'

B13



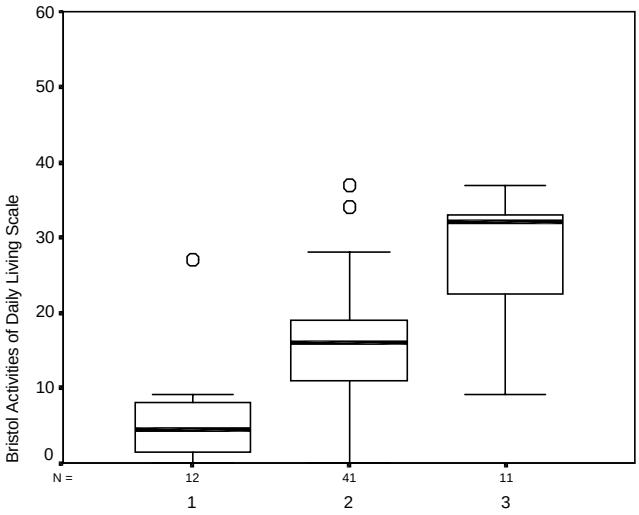
Carer proxy-rated EQ-5D 'self-care'

B14



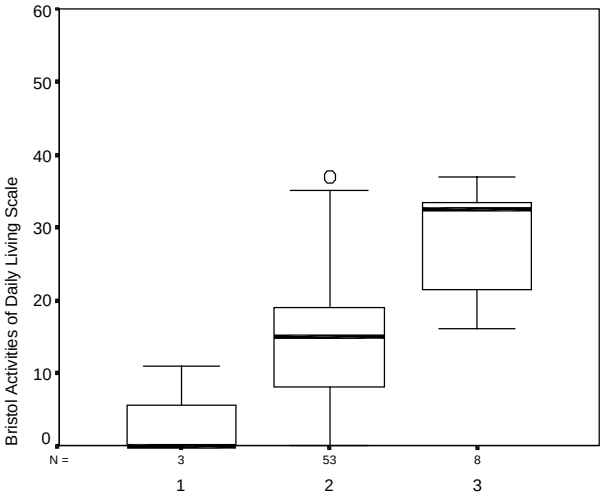
Clinician proxy-rated EQ-5D 'self-care'

B15



Carer proxy-rated EQ-5D 'usual activities'

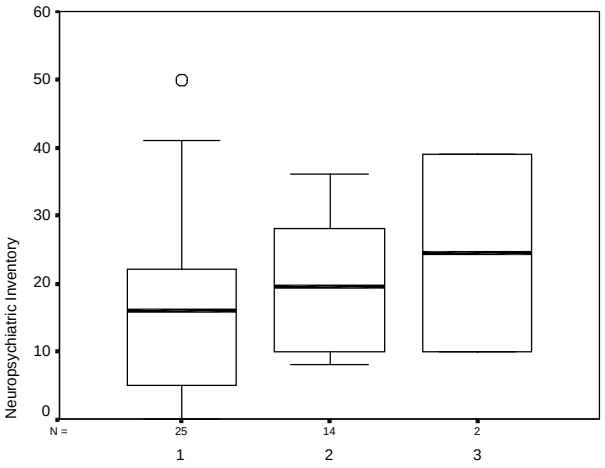
B16



Clinician proxy-rated EQ-5D 'usual activities'

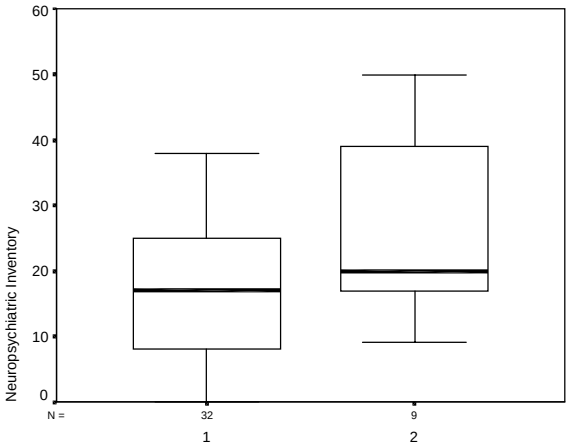
NPI scores and EQ-5D dimension comparisons

B17



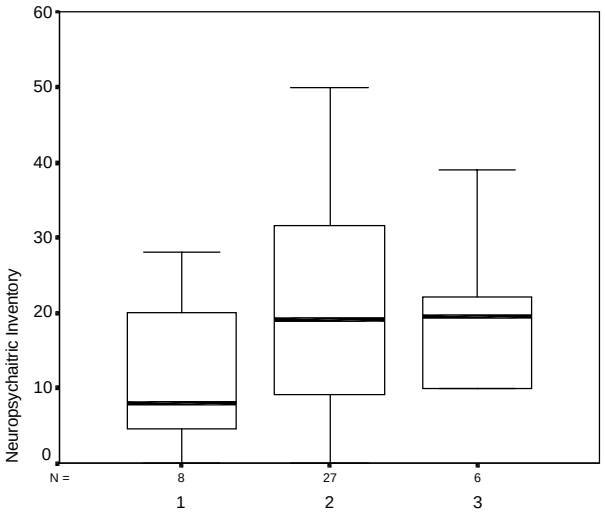
Carer proxy-rated EQ-5D 'self-care'

B18



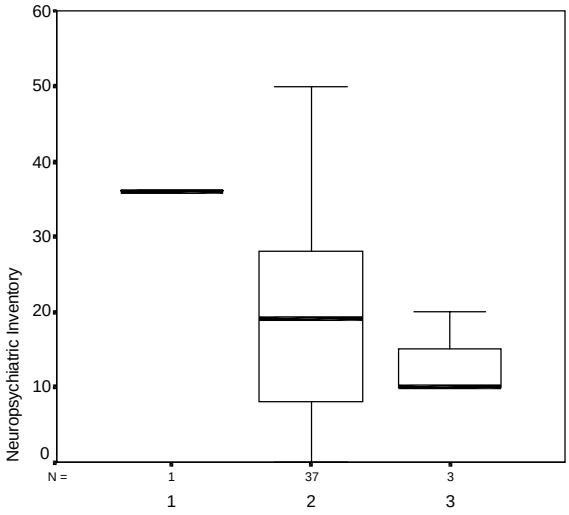
Clinican proxy-rated EQ-5D 'self-care'

B19



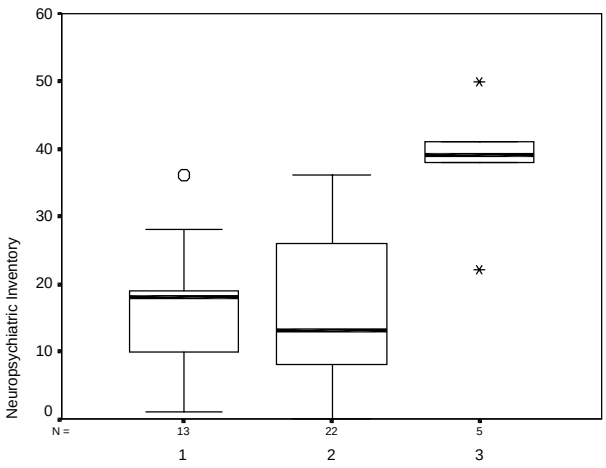
Carer proxy-rated EQ-5D 'usual activities'

B20



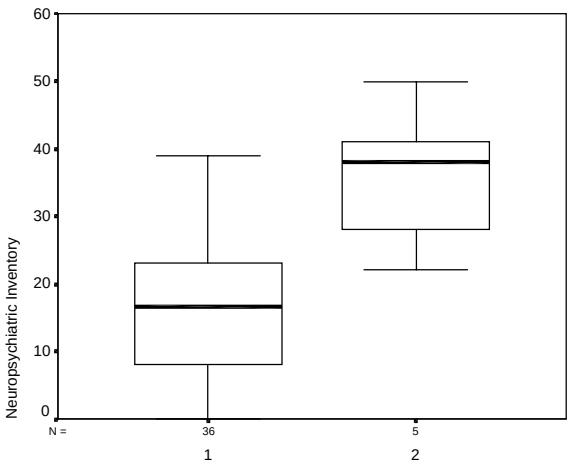
Clinician proxy-rated EQ-5D 'usual activities'

B21



Carer proxy-rated EQ-5D 'anxiety/depression'

B22



Clinician proxy-rated EQ-5D 'anxiety/depression'

*Data only available on NPI on subsample of patients (n=41). NPI maximum score is 144. Boxplots scaled to 60 as 50 was the maximum score.