Health-related quality of life in people with advanced dementia: a comparison of EQ-5D-5L and QUALID instruments

Abstract

Background: Assessing health-related quality of life (HRQOL) in people with advanced dementia is challenging but important for informed decision-making. Proxy measurement of this construct is difficult and is often rated lower than self-report. Accurate proxy rating of quality of life in dementia is related to identification of concepts important to the person themselves, as well as the sensitivity of the measures used. The main aim of this study was to compare the performance of two instruments - QUALID and EQ-5D-5L - on measuring HRQOL in people with advanced dementia.

Methods: In a sub-study nested within a cluster-RCT we collected proxy(nurse)-completed EQ-5D-5L and QUALID measures at baseline, 3, 6, 9 and 12 months’ follow-up for people with advanced dementia, residing in 20 nursing homes across Australia. Spearman’s rank correlations, partial correlations and linear regressions were used to assess the relationship between the HRQOL instrument scores and their changes over time.

Results: The mean weight from 284 people for the EQ-5D-5L and QUALID at baseline were 0.004 (95% CI -0.026, 0.033) and 24.98 (95% CI 24.13, 25.82), respectively. At 12 months’ follow-up, 115 participants remained alive. EQ-5D-5L weights and QUALID scores at baseline and at follow up were moderately correlated ($r = -0.437; p<0.001$ at 12 months). Changes within QUALID and EQ-5D-5L across the same follow up periods were also correlated ($r=-0.266; p=0.005$). The regression analyses support these findings.

Conclusion: Whilst these quality of life instruments demonstrated moderate correlation, the EQ-5D-5L does not appear to capture all aspects of quality of life that are relevant to people with advanced dementia and we cannot recommend the use of this instrument for use within this population. The QUALID appears to be a more suitable instrument for measuring HRQOL in people with severe dementia, but is not preference-based, which limits its application in economic evaluations of dementia care.

Keywords: Dementia, Health-related Quality of Life, QALY, QUALID, EQ-5D-5L, proxy, advanced dementia
1. Introduction

Health related quality of life (HRQOL) measures are used both clinically and in the decision-making process of healthcare resources allocation [1]. The challenging nature of measuring HRQOL is further emphasised by a degree of cognitive impairment, such as with dementia. It is widely recognised that the best source of data for HRQOL assessment is the person themselves. It is possible for people with mild to moderate dementia to give clear reports of their quality of life [2, 3] using dementia-specific measures [4] and through on-on-one interviews [5]. It is more difficult for people with advanced dementia to report on HRQOL with non-dementia-specific instruments such as the EQ-5D [6]. The nature of advanced dementia eliminates the possibility of getting first-hand experience from the patient, as self-reporting HRQOL requires a level of cognition and self-awareness which is unattainable by people with dementia in advanced stages of the disease [7, 8]. To address these issues, HRQOL data can be obtained by proxy, such as a relative or professional caregiver [9].

The choice of instrument for assessing HRQOL in people with dementia is important. Generic measures, such as the EQ-5D, are widely used and are often favoured by decision makers as they are quick to complete, adaptable to a large number of disease conditions and easily translated into Quality-Adjusted Life Years (QALYs) [10]. However, a European consortium on outcome measures in dementia care has concluded that the 3-level version of the EQ-5D is unsuitable for use in people with dementia and has made the recommendation that dementia-specific measures should be used, in particular when assessing the effect of psychosocial interventions for people with dementia [11]. Since this recommendation, the 5-level EQ-5D version has been introduced for use in dementia [12], with the aim of improving the instrument’s sensitivity and reducing the ceiling and floor effects [13]. However, a study on inclusion of a cognitive dimension in the EQ-5D for evaluating HRQOL in people with cognitive impairment found no added benefit of changing the instrument [14].

Dementia-specific instruments potentially capture broader and more relevant aspects of HRQOL in people with dementia, compared to generic measures, such as the EQ-5D [10]. At least 16 dementia-specific instruments exist for measuring HRQOL [15-17]. These differ in breadth and focus, although some common features have been reported: social relations or interaction, self-esteem, and mood [10]. Dementia-specific instruments also target different stages of dementia, and some have been specifically designed for self- or proxy-completion [4, 18]. However, the relationship between dementia-specific measures and generic preference-based QOL measures is not always clear, nor is the usefulness of dementia-specific instruments in economic evaluations.

We investigated the correlations between the two instruments and assessed the sensitivity of the Quality of Life in Late-stage Dementia (QUALID) to changes in EQ-5D-5L weights in the sub-study. In this paper we compare the performance of proxy-completed QUALID and EQ-5D-5L instruments for measuring HRQOL in people with advanced dementia. To the best of our knowledge, there has not been a study comparing the performance of QUALID and EQ-5D-5L instruments for people with advanced dementia thus far.

2. Methods

2.1 Data

This is a secondary analysis of data obtained from a sub-study conducted within the IDEAL trial, conducted as a pragmatic parallel cluster randomised controlled trial investigating the effectiveness of facilitated family case conferencing compared to usual care in 20 Australian nursing homes [12, 19]. Participants included in the trial had a Functional Assessment Staging Tool (FAST) [7] in dementia score of 6a or higher and an Australia-modified Karnofsky Performance Status (AKPS) [20] score of ≤80. These criteria were chosen because a FAST stage 7c combined with functional dependency (measured here by the AKPS) is predictive of an average survival of <6 months, and the study’s primary endpoint focused on end of life care [21]. The FAST tool was initially designed for assessment of
disease severity in people with Alzheimer’s disease, it has been used to assess severity in people with broader dementia diagnoses in other studies (e.g. - [22, 23]). The detailed inclusion criteria are outlined in the protocol paper [19].

In the intervention arm, registered nurses were trained as Palliative Care Planning Coordinators (PCPCs). The PCPCs’ role was to identify optimal time-points for case conferencing, as well as to organise case conferences between the person with dementia, their family, multi-disciplinary nursing home staff and community health professionals involved in their care. The primary endpoint was symptom management, comfort and satisfaction with care at the end of life, rated with the End of Life in Dementia (EOLD) Scales [24]. Secondary outcomes included the person’s HRQOL (via the EQ-5D-5L and QUALID), resource use, and staff attitudes and knowledge of dementia care. Eligible participants were followed up every 3 months for 12 months. Recruitment took place between February 2013 and August 2015.

We found no difference in the EOLD scale or in QUALID in the IDEAL trial [12]; therefore, for the purposes of this analysis, participants from both arms were combined (n=284). Proxy HRQOL was measured by specially-trained nurses familiar with participants, using the EQ-5D-5L and QUALID instruments at baseline, three months, six months, nine months and 12 months [12]. Measures for each participant were completed by the same nurse at each time point, but may have been completed by different nurses across the study timeline.

Ethics approval was granted by the Human Research Ethics Committee (HREC) of the University of New South Wales, Australia and ratified by HRECs at the University of Technology Sydney and Queensland University of Technology, Australia.

2.2 EQ-5D-5L and QUALID instruments
The EQ-5D-5L is a five level, five-item version of the world’s most widely used multi-attribute utility instrument EQ-5D [6], which has been revised to include a larger number of severity levels among its response options (therefore giving greater sensitivity). The EQ-5D-5L assesses HRQOL across five dimensions: mobility, self-care, usual activities, pain/discomfort and anxiety [13]. The respondent (or proxy) rates each domain on a 5-point Likert scale from ‘no problems with a task’ to ‘unable to perform a task’ [13]. Raw EQ-5D-5L utilities were weighted with Australian-specific weights [25], where 1 represents full health, 0 represents death, and negative weights represent states worse than death.

The 11-item QUALID scale is an advanced dementia-specific HRQOL measure that has been specifically designed for proxy rating by caregivers. QUALID focuses on the HRQOL of the person over the span of the last 7 days by asking the proxy to rank 11 statements on a five-point scale. Possible scores range from 11 to 55, with lower scores representing the highest quality of life [26]. Total QUALID scores were calculated by summing the proxy responses to each of the 11 questions. The Quality of Life in Late-stage Dementia (QUALID) scale was developed for assessing QOL in people with advanced dementia who may be unable to communicate coherently on discrete aspects of QOL. It specifically focuses on measuring aspects that are more relevant to the person with dementia, rather than other more widely recognised contributors to QOL in the general population [26]. QUALID was designed to be administered by a caregiver or family member of a person with advanced dementia residing in long-term care facilities [26]. It has been reported to have strong internal consistency, with high test-retest reliability [15, 26]. However, it is has not been found to be correlated with measures of cognition, or activities of daily living [15, 26]. It is unclear how QUALID compares to generic QOL measures, and whether it is suitable for economic evaluation

The QUALID was selected due to being designed purposefully for proxy-rating and its high internal validity, as specified in the protocol paper [19]. The EQ-5D-5L was chosen for comparison with the QUALID due to the potential
to undertake economic evaluation, as the QUALID cannot be used to derive Quality-Adjusted Life-Years (QALYs), and also in order to validate the EQ-5D-5L instrument in this population using proxy rating.

2.3 Statistical analyses
Since assessment focused on changes in the participant’s HRQOL over time with exposure to the case conference intervention, we used Spearman’s rank correlation to assess participant HRQOL and compared QUALID and EQ-5D-5L weights at each follow-up interval. We compared 3-month changes (from baseline to 6 to 9 to 12 months) in QUALID and EQ-5D-5L using partial correlation, controlling for age, gender and baseline EQ-5D-5L and QUALID measurements. We considered correlations to be strong with scores of 0.5 and higher, and moderate with scores of 0.25-0.49 [27]. Correlations less than and equal to 0.24 were considered weak [27].

We used linear regressions to investigate the relationship between changes in EQ-5D-5L and QUALID over three-month intervals, and controlled for age, gender, as well as the EQ-5D-5L and QUALID scores at the start of each interval being assessed. All statistical analysis was carried out using STATA version 14 [28].

3. Results
Of the 284 participants with dementia 179 (63%) were female and 105 (37%) male (Table 1). Forty-five percent were widowed, and 43% were married, with the rest divorced. A third (33%) had primary school certificate as their highest level of education attainment; a further 35% and 15% had completed high school and higher school certificates, respectively. Thirteen percent had some form of diploma or tertiary qualification, and 3% had no formal education.

The participants’ FAST (level of function) scores ranged between levels 6a and 7f, with 50% of the study population being in groups 7c and d, corresponding to ‘Ambulatory ability lost (cannot walk without personal assistance)’ and ‘Ability to sit up without assistance lost (e.g., the individual will fall over if there are no lateral rests [arms] on the chair)’, respectively. Less than half the study population (115 of 284) remained alive at the 12-month follow-up (Table 2). No statistically significant differences in demographics, survival or HRQOL outcomes between control and intervention groups were observed (data not presented) [12].

The mean QUALID score at baseline was 24.98 (SD=7.23), whilst the mean EQ-5D-5L weight was 0.004 (SD=0.25) (Table 2). The mean QUALID score declined slightly to 24.23 (SD=6.14) at 12 months, but the difference was not statistically significant. The EQ-5D-5L weights also experienced a consistent and statistically insignificant reduction with a mean of -0.045 (SD=0.22) at 12-month follow-up.

For participants remaining at the final follow up (n=115), the mean QUALID score increased from 24.23 (SD=6.14) at baseline to 25.43 (SD=7.45) at 12 months, although the increase was not statistically significant (p=0.104). Similarly, the change in the EQ weights from -0.0026 (SD=0.24) at baseline to -0.045 (0.22) at 12 months was not statistically significant either (p=0.083). For participants who died before 12 months’ follow up, mean baseline scores were 0.008 (SD=0.25) and 24.67 (SD=7.07) for EQ-5D-5L and the QUALID, respectively.

Table 1: Demographic characteristics and FAST scores of study population at baseline and 12 months follow up

<table>
<thead>
<tr>
<th></th>
<th>At Baseline</th>
<th>Surviving to 12 months</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean age at enrolment (mean, n (SD))</td>
<td>85 (8) 284</td>
<td>85 (9) 115</td>
</tr>
<tr>
<td>Gender, n (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>179 63%</td>
<td>79 69%</td>
</tr>
<tr>
<td>Male</td>
<td>105 37%</td>
<td>36 31%</td>
</tr>
<tr>
<td>Marital status, n (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Married</td>
<td>121 43%</td>
<td>48 42%</td>
</tr>
<tr>
<td>Separated</td>
<td>3 1%</td>
<td>2 2%</td>
</tr>
</tbody>
</table>
The correlations between QUALID and EQ-5D-5L were moderate for each time point (Table 3). Correlations were all highly statistically significant ($p < 0.001$) and ranged between -0.3 at 9 months and -0.44 at 12 months. Intraclass correlation coefficients for residential facilities, resulting from two-way mixed effects model regression, were also very low, ranging between 0.00 and 0.05.
In order to investigate the nature of the relationship between QUALID and EQ-5D-5L, we correlated the scores/utilities for each of the five domains of EQ-5D-5L with QUALID at baseline (Table 6). The anxiety/depression and pain/discomfort domain scores were moderately correlated with QUALID scores at baseline ($r=0.46$, $p<0.001$; $r=0.30$, $p<0.001$, respectively). Baseline QUALID scores were weakly correlated with the ‘self-care’ ($r=0.12$, $p=0.043$) and ‘usual activities’ ($r=0.13$, $p=0.027$) domains of the EQ-5D-5L, while ‘mobility’ did not appear to have any
correlation with QUALID scores. Bland-Altman plots were also constructed to explore the convergence between the two instruments, but did not reveal any additional information.

| Table 6: Correlations between QUALID and EQ-5D-5L health state dimensions at baseline |
|--------------------------------------|------------------|--------|--------|
| EQ-5D-5L domains                    | Spearman’s rank correlation coefficient | p      | n      |
| Mobility                             | 0.047            | 0.433  | 283    |
| Self-care                            | 0.120            | 0.043  | 283    |
| Usual activities                     | 0.131            | 0.027  | 283    |
| Pain/discomfort                      | 0.299            | <0.001 | 280    |
| Anxiety/depression                   | 0.455            | <0.001 | 281    |

4. Discussion

Both the QUALID scores and EQ-5D-5L weights increased from baseline to 12-month follow-up in study participants with exposure to the case conference intervention, although the increases were not statistically significant. The main interest of this sub-study, however, was how these dementia-specific quality of life scales performed in relation to each other. The results indicate a consistent moderate correlation between the proxy-completed EQ-5D-5L and the QUALID across five different time points over a 12-month follow-up period (Table 3, Appendix Figure 1). The strength of the correlations between the two instruments appears to be on par with other instances of comparisons of preference-based measures with non-preference-based ones [29].

Changes within each measure were also moderately correlated across these time points (Tables 4 – 5). This indicates further consistency in the correlation between the two instruments. This relationship, unsurprisingly, occurs between the QUALID scores and the ‘pain/discomfort’ and ‘anxiety/depression’ domains of EQ-5D-5L, and, to a lesser extent, the association with the ‘self-care’ and ‘usual activities’ domains (Table 6). The QUALID does not measure mobility, thus explaining the lack of correlation with that domain of the EQ-5D-5L. Overall, there is insufficient evidence to support using the EQ-5D-5L as a proxy instrument for this population. Previous research found the EQ-5D-5L to be suitable for use in people with mild dementia [14, 30, 31]. In this sub-study, however, the EQ-5D-5L weights were very close to zero, indicating that it was not sensitive to changes in HRQOL in study participants in the advanced stage of dementia (Table 2). The majority of participants had significant cognitive impairment, and many were in palliative stages of care (as suggested by a 60% mortality rate over the year-long follow up), which possibly made it difficult for proxies to rate HRQOL based on EQ-5D-5L constructs.

QUALID provided a broader scope for proxy assessment of HRQOL, using constructs more relevant to advanced and end-stage dementia such as comfort and happiness [26]. Nevertheless, QUALID is subject to the usual drawbacks of a disease-specific measure of HRQOL, in that it does not reflect preferences for health states. It is, therefore, not possible to obtain utility values with QUALID, which restricts cross-study comparisons and decision making on investing in dementia-relevant services. As well, QUALID has not proven superior, or any more reliable, than other dementia-specific quality of life scales [2, 32]

4.1 Self and proxy-completion

Proxy measurement of intangible concepts like QOL and HRQOL are difficult when the person with dementia has deteriorated to such the level that they have lost the capacity to give expression to most indicators being assessed. The reason that proxy rating of QOL is sought from others in regular contact with the person, is that the person may have lost the capacity to provide a subjective response to how they feel about their life, and may not be able to evaluate their current life quality by weighing up both positive and negative experiences [33-36]. For people with
dementia who are able to provide information on their QOL, they often provide higher ratings than the proxy responders, unless proxies have a very good appreciation of what the person’s current aspirations and experiences are [33-36].

To obtain information on aspects of HRQOL that are meaningful to the person with dementia, it is important that the proxy informant is able to make unbiased judgements on the concepts being assessed [37]. Proxy HRQOL rating bias is related to personal HRQOL, financial situation and age [38]. Differences have also been found in the proxy HRQOL ratings of informal and formal dementia caregivers, since it is necessary for the proxy to reflect on the meanings of QOL concepts in the dementia context, which some caregivers find difficult to do [37, 39].

Proxy ratings require detailed attention to the process of dementia and its effects on peoples’ identities, self-perceptions, capacities and value preferences. Another issue is the time-factor, i.e. assessment of HRQOL at a given time or over longer periods, whereby proxies are asked to make judgments on goodness of life as a whole, rather than on goodness in relation to a particular period of time [40]. This suggests that proxy rating of HRQOL in a person with dementia is a factor of proxy insight of what HRQOL means for the person with dementia, the strength and quality of their relationship with the person, and the constancy of their presence with the person in the period that measurement occurs.

In this sub-study, proxy measurement of HRQOL was provided by formal caregivers who had close associations with individual participants with dementia, which may have helped to avoid some of the potential bias with reliance on informal caregivers who spent less time with the participant during the assessment period. The analyses investigated the intraclass correlations of different residential facilities, demonstrating very low or no correlations between individuals, suggesting poor reliability between raters, [41] or a potential lack of variability among raters, as evidenced by little variation in the data.

### 4.2 QUALID and preference based measures for dementia

To progress QOL assessment in dementia it would be valuable to have access to a preference-based measure based on QUALID. Rowen et al. [42] have made such an attempt with the development of DEMQOL-U and DEMQOL-proxy-U on the basis of DEMQOL and DEMQOL-proxy scales. The DEMQOL-proxy-U is limited to four domains: positive emotion, memory, appearance and negative emotion, each with four levels[42]. These are relatively narrow in scope, and omit any function or physical wellbeing sub-scales. Furthermore, there is some evidence suggesting that DEMQOL-proxy and DEMQOL-proxy-U are more responsive to changes in depression and delirium symptoms rather than physical symptoms [34, 43]. While vastly different from EQ-5D-5L, it could be argued that DEMQOL-proxy-U may yield similar scores to EQ-5D-5L in populations with dementia as advanced as in this study - i.e. close to zero. Development of a preference-based measure using HRQOL concepts contained in QUALID could potentially address this issue.

### 4.3 Strengths and limitations

One of the strengths of this sub-study was the use of two validated QOL scales in measuring HRQOL in people with advanced dementia until their death, as rated by proxies. The QOL measurements were taken at five different time points, further adding to the strengths and robustness of the dataset. The quality and reliability of the data across multiple time points helped to track and compare changes in QOL for participants, thereby providing novel insights which may have not occurred with reliance on either the QUALID or the EQ-5D-5. These data provide rich information on HRQOL issues that will help to improve end-of-life care and experiences for people with advanced dementia, and will contribute to the discussion on how best to measure health related quality of life in people with significantly reduced cognitive abilities. EQ-5D-5L is an extension of a commonly used HRQOL measure EQ-5D-3L.
which, at the time this study was conducted, remained to be validated in a number of clinical settings such as advanced dementia, and therefore, provides a reliable point for comparison of a disease specific measure such as QUALID.

It is difficult to establish whether the current study is subject to proxy bias. Because this study utilised nursing home workers as proxies, and not informal carers or relatives, it is possible that the bias was avoided or, at least, minimised. It is also possible that nurses may introduce the bias of perceived benefit of care delivered by them [44]. Another concern is the risk of inconsistencies in proxy-HRQOL measures, as these have been completed by different nurses at different study timelines. As the same nurse completed both instruments at the same time point for each participant, the consistency between the two instruments should not be affected. The longitudinal validity of the study may be reduced by the fact that different raters completed the instruments at different timepoints, potentially limiting the comparability of the instrument findings over the study period. However, the experience of nurses conducting the rating, and their familiarity with each participant may have had an effect on the measurements/weights reported [9]. While the validity of the QUALID has been demonstrated [26, 45], it remains a subjective measure (e.g. having to judge whether or not a subject enjoys eating food) and the final scores may be rater-biased.

5. Conclusions

One of the ways to develop and evaluate interventions is to assess these concepts using valid and reliable dementia-specific QOL and HRQOL scales. This sub-study employed the QUALID and the EQ-5D-5L to obtain proxy ratings of HRQOL in people with advanced dementia living in residential care homes over 12 months. The study showed that QUALID is a suitable and reliable instrument for proxy measurement of HRQOL in people with severe dementia, compared to EQ-5D-5L, as it is more sensitive to the particular features of HRQOL in dementia. The EQ-5D-5L fails to identify aspects of HRQOL that are obtained with the QUALID. The main limitation of QUALID is that it is not preference-based and cannot easily be used in economic evaluations, where QALYs are the main outcome. Further research should focus on further comparisons of QUALID with other generic and disease-specific HRQOL measures and development of a preference-based measure based on QUALID.
Compliance with Ethical Standards:

Funding:
This study was funded by the Australian Department of Health (previously Department of Health and Ageing) (http://www.health.gov.au/).

Conflict of Interest:
- Elizaveta Sopina declares that she has no conflict of interest
- Lynn Chenoweth declares that she has no conflict of interest
- Tim Luckett declares that he has no conflict of interest
- Meera Agar declares that she has no conflict of interest
- Georgina M Luscombe declares that she has no conflict of interest
- Patricia M. Davidson declares that she has no conflict of interest
- Constance D Pond declares that she has been on an advisory committee on dementia for Eli-Lilly, has received speaker fees for speaking about dementia to Australian General Practitioners, Primary Health Networks and Alzheimer’s Australia, as well as airfares and accommodation to prepare educational material on dementia for General Practitioners
- Jane Phillips declares that she has no conflict of interest
- Stephen Goodall declares that he has no conflict of interest

Ethical approval:
This article does not contain any studies with animals performed by any of the authors.

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. Human Research Ethics Committees of the University of NSW, University of Technology Sydney, Queensland University of Technology approved the trial. The trial was registered under trial ID no ACTRN12612001164886 https://www.anzctr.org.au/Trial/Registration/TrialReview.aspx?id=362912

Informed consent:
Informed consent was obtained from family members with legal authority on behalf of all study participants.

References
28. StataCorp, Stata Statistical Software: Release 14. 2015, College Station, TX: StataCorp LP.