Psychometric properties and feasibility of use of dementia specific quality of life instruments for use in care settings: a systematic review

Laura J. Hughes,¹ Nicolas Farina,¹ Thomas E. Page,² Naji Tabet,¹ and Sube Banerjee¹

¹Centre for Dementia Studies, Brighton and Sussex Medical School, Brighton, UK
²School of Psychology, University of Kent, Canterbury, UK

ABSTRACT

Background: Over 400,000 people live in care home settings in the UK. One way of understanding and improving the quality of care provided is by measuring and understanding the quality of life (QoL) of those living in care homes. This review aimed to identify and examine the psychometric properties including feasibility of use of dementia-specific QoL measures developed or validated for use in care settings.

Design: Systematic review.

Methods: Instruments were identified using four electronic databases (PubMed, PsycINFO, Web of Science, and CINAHL) and lateral search techniques. Searches were conducted in January 2017. Studies which reported on the development and/or validation of dementia specific QoL instruments for use in care settings written in English were eligible for inclusion. The methodological quality of the studies was assessed using the COSMIN checklist. Feasibility was assessed using a checklist developed specifically for the review.

Results: Six hundred and sixteen articles were identified in the initial search. After de-duplication, screening and further lateral searches were performed, 25 studies reporting on 9 dementia-specific QoL instruments for use in care home settings were included in the review. Limited evidence was available on the psychometric properties of many instruments identified. Higher-quality instruments were not easily accessible or had low feasibility of use.

Conclusions: Few high-quality instruments of QoL validated for use in care home settings are readily or freely available. This review highlights the need to develop a well-validated measure of QoL for use within care homes that is also feasible and accessible.

Key words: dementia, quality of life, care settings, feasibility, psychometrics

Introduction

There are approximately 16,000 care homes in the UK providing care for an estimated 416,000 people (Care Quality Commission, 2017a; NIHR, 2016); in the US over 2 million individuals are cared for in approximately 45,000 facilities (Centre for Disease Control and Prevention, 2017). These numbers are expected to increase with an ageing population (Prince et al., 2014). It is estimated that anywhere between 40% and 86% of people living in care homes have a dementia, either diagnosed or undiagnosed, or some form of memory impairment (Jagger and Lindesay, 1997; Livingston et al., 2017). The large numbers of people living in care homes has led to a need to understand the outcomes and experiences of care by residents in order to make improvements to care. The quality of care provided in care homes and its consistency has been questioned, with calls to increase the services provided to vulnerable adults (Care Quality Commission, 2017b; Department of Health, Prime Minister’s Office, 2015). It has been argued that one way we may be able to better understand the outcomes of care in order to improve it is through the measurement and understanding of quality of life (QoL) (Black, 2013; Edelman et al., 2005).
Due to a growing recognition of the importance of the outcomes of care as well as the process of care (Sloane et al., 2005), there has been a growing emphasis on measuring QoL as a means of understanding and improving care in care settings. An increasing amount of research in care homes is being carried out with QoL as an important outcome for evaluating interventions (Aspden et al., 2014). A number of dementia-specific QoL measures have been developed; however, most of these were developed and evaluated in community-dwelling populations, rather than in care homes (Bowling et al., 2015).

Previous reviews have examined dementia-specific QoL instruments in general (Bowling et al., 2015; Ready and Ott, 2003), and generic QoL measures available for use in care homes (Aspden et al., 2014). However, generic QoL measures have limitations in capturing the experiences of people with dementia (Smith et al., 2005). No previous review has focused on the usability of the instruments (e.g. availability, cost, training, and how easy is it to score and interpret) in care homes settings. Ultimately, the usability of the instrument is likely to dictate whether an instrument is used or not. We therefore aimed to add to previous reviews by carrying out a systematic review to identify and examine not only the psychometric properties but also the usability of disease-specific instruments that measure the QoL of people living in care homes.

**Methods**

The protocol for this review is registered in the International Prospective Register of Systematic Reviews (PROSPERO) – CRD42017046272.

**Inclusion criteria**

Studies were included in the review if they met the following criteria: 1) They described the development and/or evaluation of an instrument, or described the adaptation and evaluation of an existing instrument for a care home population; 2) the instrument was a dementia-specific QoL instrument, and studies evaluating only generic health-related QoL instruments in a care home population were excluded; 3) the study population included residents living in a care home (including nursing homes), and data on this group were presented separately from any others studied; and 4) studies were published in English. There were no exclusion criteria based on diagnosis, or non-diagnosis, of dementia.

**Search strategy**

Articles were identified from initial searches in four electronic databases: PubMed, PsycINFO, Web of Science, and the Cumulative Index to Nursing and Allied Health Literature (CINAHL). All searches were conducted in January 2017. There were no restrictions on date of publication. The following four combined search terms were used: 1) quality of life OR QOL OR health related quality of life OR HRQOL OR HRQL; AND 2) dementia OR Alzheimer’s; AND 3) residential facilities OR residential care institutions OR long-term care OR nursing homes OR care homes OR residential care homes, AND 4) measure development OR valid* OR reliab* OR accuracy OR feasibility OR scale. Lateral searches involved checking the references of included studies (snowballing), and further searching for identified measures on PubMed and Google search engine. Two independent reviewers (LH & NF) screened article titles and abstracts against the predefined inclusion criteria. Full text articles were sought for all relevant studies. Any disagreements regarding inclusion were resolved through discussion by the two reviewers.

**Data extraction**

Two reviewers (LH & TEP) independently extracted the following data from included full-texts: name of instrument, country, language of instrument, sample characteristics (i.e., age, gender, and dementia severity), study design, measurement domains, individual items, number of items, response format, and evidence of reliability and validity.

The reviewers also extracted the following data about the usability of each instrument through reading the identified articles, internet searches (Google.com, first 30 hits) and a measure specific database (Mapi; https://mapi-trust.org/). Data included:

- Existence of a dedicated website (Yes/No) – The accessibility of an online website or repository, which contains dedicated information about the instrument
- User guide available (Yes/No) – The availability of a user guide or instructions about how to use the instrument
- Cost (paid, free, or other) – The cost of using the questionnaire, broken down by user type when provided
- Training – Details about what training is available for the instrument, whether it is a prerequisite, and the costs involved
- Time to complete measure (minutes)
- Time period the instrument captures
- Need for specialist software for scoring and interpretation (Yes/No)
- Guidance available for scoring and interpretation (Yes/No)

“Not known” was used in instances where we were unable to identify measure specific information.
Quality assessment
The methodological quality of the studies was assessed using the Consensus-based Standards for the selection of health Measurement Instruments (COSMIN) checklist (Mokkink et al., 2010). This is a standardized tool which assesses the measurement properties of health-related instruments across nine domains (internal consistency, reliability, measurement error, content validity [including face validity], construct validity [subdivided into structural validity, hypotheses testing, and cross-cultural validity], criterion validity and responsiveness) with each domain rated using 5–18 items. Each item is rated as “excellent”, “good”, “fair”, or “poor” quality. A methodological quality score for each measurement property is obtained by taking the lowest rating of any item in that box (“worst score counts”). Two independent reviewers (LH & TEP) assessed the methodological quality of the included studies using the checklist. Studies will not be excluded based on COSMIN scores. Any disagreements in scoring were resolved through discussion and advice from a third reviewer (NF).

Data synthesis
A narrative synthesis was adopted to assess the feasibility of the QoL measures for use in care homes. No formal frameworks or criteria exist for the assessment of the feasibility of the QoL measures for use in care homes. A narrative synthesis was adopted to assess the methodological quality of the included studies. The measures were: the Dementia Quality of Life (DQoL) (Brod et al., 1999) (n = 4), Quality of Life in Alzheimer’s Disease (QOL-AD) (Logsdon et al., 1999) (n = 1), Quality of Life in Alzheimer’s Disease nursing home version (QOL-AD NH) (adapted by Edelman and Fulton, unpublished work) (n = 6), QUALIDEM (Ettema et al., 2007b) (n = 7), Quality of Life in Late-Stage Dementia (QUALID) (n = 6) (Weiner et al., 1999), Dementia Care Mapping (DCM) (Bradford Dementia Group, 2005) (n = 5), Alzheimer Disease Related Quality of Life (ADRQL) (Rabins et al., 1999) (n = 3), ADRQL revised (Kasper et al., 2009) (n = 1), and the Quality of Life in Dementia (QOL-D) (Albert et al., 1996) (n = 1).

Study and measure characteristics
The sample size of the studies ranged from 20 (Beer et al., 2009) to 973 (Bouman et al., 2011). Table 1 outlines the characteristics of included studies. The studies recruited participants from the United Kingdom (n = 8), the United States (n = 2) Netherlands (n = 3), Spain (n = 4), Germany (n = 5), Sweden (n = 1), and Norway (n = 2).

Of the nine instruments identified, three instruments contained self-report and a proxy report questions (QOL-AD, QOL-AD NH, and QOL-D), one instrument contained a self-report questions only (DQoL), four instruments contained a proxy report questions only (QUALIDEM, QUALID, ADRQL, and ADRQL revised), and one instrument consists of a proxy observation tool (DCM). The number of measurement domains ranged from 1 (QOL-AD, QOL-AD NH, and QUALID) to 9 (QUALIDEM 37 item version). Five instruments have a Likert scale response format (DQoL, QOL-AD, QOL-AD NH, QUALID, and DCM), two have a frequency scale format (e.g., never, sometimes, often) (QUALIDEM and QOL-D), and two instruments have a dichotomous response format (e.g., agree/disagree) (ADRQL and ADRQL revised). The number of items in each instrument ranged from 6 (affect subscale of QOL-D) to 40 (ADRQL revised). See Supplementary Table 1 for full details.

Six instruments were developed specifically for use in care settings (QUALID, QUALIDEM, DCM, ADRQL, ADRQL revised, and QOL-D), and one was adapted for use in care settings and evaluated (QOL-AD NH). The QOL-AD and DQoL were evaluated for use in care settings but were not developed or adapted for this purpose.

Psychometric properties
The instrument with the most comprehensive evaluation of psychometric properties across the nine domains of the COSMIN checklist was the QUALID with seven domains assessed in two separate studies. The DQoL, QOL-AD, QUALIDEM, and DCM had four domains assessed in at least one study.
The instruments with the least domains assessed were QOL-AD NH ($n = 3$), QOL-D ($n = 3$), ADRQL ($n = 3$), and ADRQL revised ($n = 1$). See Table 2 for results of the COSMIN checklist.

**Internal consistency**

All nine instruments had internal consistency reported in at least one study. Instruments that were reported on the most were QUALID (six studies with poor to good ratings), and QUALIDEM (five studies, poor to excellent quality ratings). Most of the instruments had acceptable to good internal consistency across some studies. QOL-AD, ADRQL Revised, and QOL-AD NH all had the highest internal consistency scores in a single or small number of studies (all >0.80). QUALID had the most consistent high internal consistency scores for multiple studies (all >0.70). All other instruments had variability in scores from poor to excellent internal consistency.

**Reliability**

Test retest and inter-rater reliability were rated in eight instruments ($n = 15$). The most assessed instruments were QUALIDEM, QUALID, and DCM ($n = 4$). QUALIDEM was rated fair to good, QUALID and DCM were both rated poor to fair. QOL-AD NH was assessed in three studies (poor to fair), DQoL was assessed in two (fair), and the QOL-AD (poor) ADRQL (fair) and QOL-D (fair) were assessed in one study. Time between tests for test-retest reliability ranged from 2–3 days (QUALID) to 12 months (DQoL). There was large variability in reliability scores for instruments. QOL-D and ADRQL both had the highest reliability, from a single study each (>0.95). QUALID had the highest...
Table 1. Characteristics of included studies

<table>
<thead>
<tr>
<th>INSTRUMENT USED</th>
<th>COUNTRY</th>
<th>N</th>
<th>GENDER (% F)</th>
<th>MEAN AGE (SD)</th>
<th>COGNITIVE IMPAIRMENT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adler and Resnick (2010)</td>
<td>DQOL</td>
<td>United Kingdom</td>
<td>486</td>
<td>80</td>
<td>83.8 (8.2)</td>
</tr>
<tr>
<td>Beer et al. (2009)</td>
<td>QOL-AD NH</td>
<td>United Kingdom</td>
<td>20</td>
<td>70</td>
<td>87.5 (8.1)</td>
</tr>
<tr>
<td>Bouman et al. (2011)</td>
<td>QUALIDEM</td>
<td>Netherlands</td>
<td>973</td>
<td>71–76 across four samples</td>
<td>83.9 (5.8)</td>
</tr>
<tr>
<td>Clare et al. (2014a)</td>
<td>QUALID</td>
<td>United Kingdom</td>
<td>105</td>
<td>69</td>
<td>81.47 (8.6)</td>
</tr>
<tr>
<td>Crespo et al. (2011)</td>
<td>QOL-AD NH</td>
<td>Spain</td>
<td>209</td>
<td>79.4</td>
<td>86.07 (6.46)</td>
</tr>
<tr>
<td>Dichter et al. (2011)</td>
<td>QUALIDEM</td>
<td>Germany</td>
<td>486</td>
<td>74 mild-moderate</td>
<td>86.07 (6.46)</td>
</tr>
<tr>
<td>Dichter et al. (2013)</td>
<td>QUALIDEM</td>
<td>Germany</td>
<td>634</td>
<td>78.6 mild-severe</td>
<td>85.8 (8.0) mild-severe</td>
</tr>
<tr>
<td>Dichter et al. (2014)</td>
<td>QUALIDEM</td>
<td>Germany</td>
<td>161</td>
<td>91 mild-severe</td>
<td>85.2 (7.0) mild-severe</td>
</tr>
<tr>
<td>Dichter et al. (2016a)</td>
<td>QUALIDEM</td>
<td>Germany</td>
<td>91</td>
<td>78 mild-severe</td>
<td>84.7 (8.3) mild-severe</td>
</tr>
<tr>
<td>Dichter et al. (2016b)</td>
<td>QOL-AD NH</td>
<td>Germany</td>
<td>29</td>
<td>66</td>
<td>88.2 (7.1)</td>
</tr>
<tr>
<td>Edelman et al. (2005)</td>
<td>QOL-AD NH</td>
<td>United Kingdom</td>
<td>172</td>
<td>83.7</td>
<td>85.8 (6.4)</td>
</tr>
<tr>
<td>Ettema et al. (2007a)</td>
<td>QUALIDEM</td>
<td>Netherlands</td>
<td>202</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Ettema et al. (2007b)</td>
<td>QUALIDEM</td>
<td>Netherlands</td>
<td>238</td>
<td>76.5</td>
<td>84.69 (6.85)</td>
</tr>
<tr>
<td>Falk et al. (2007)</td>
<td>QUALID</td>
<td>Sweden</td>
<td>169</td>
<td>88</td>
<td>86 (6.0)</td>
</tr>
<tr>
<td>Fossey et al. (2002)</td>
<td>DCM</td>
<td>United Kingdom</td>
<td>123 (A)</td>
<td>76 cohort A</td>
<td>82.3 (7.4) cohort A</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>54 (B)</td>
<td>35 cohort B</td>
<td>78.2 (9.1) cohort B</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Garre-Olmo et al. (2010)</td>
<td>QUALID</td>
<td>Spain</td>
<td>160</td>
<td>76.9</td>
<td>82.6 (11.6)</td>
</tr>
<tr>
<td>Kasper et al. (2009)</td>
<td>ADRQL</td>
<td>United States</td>
<td>89</td>
<td>–</td>
<td>– (–)</td>
</tr>
<tr>
<td>León-Salas et al. (2011)</td>
<td>QOL-AD</td>
<td>Spain</td>
<td>101</td>
<td>88.1</td>
<td>83.2 (6.3)</td>
</tr>
</tbody>
</table>
Table 1. Continued

<table>
<thead>
<tr>
<th>INSTRUMENT USED</th>
<th>COUNTRY</th>
<th>N</th>
<th>GENDER (% F)</th>
<th>MEAN AGE (SD)</th>
<th>COGNITIVE IMPAIRMENT MEAN (SD)/N (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mjorud et al. (2014)</td>
<td>QUALID</td>
<td>Norway</td>
<td>661</td>
<td>71.4</td>
<td>85.3 (8.6)</td>
</tr>
<tr>
<td>Moyle et al. (2012)</td>
<td>DQoL, QOL-AD NH</td>
<td>United Kingdom</td>
<td>61</td>
<td>70.5</td>
<td>10–19 = 22 (36.1%), 0–9 = 1 (1.6%)</td>
</tr>
<tr>
<td>Røen et al. (2015)</td>
<td>QUALID</td>
<td>Norway</td>
<td>169</td>
<td>56.6</td>
<td>69.2 reliability study 83.2 (7.5) reliability study</td>
</tr>
<tr>
<td>Sloane et al. (2005)</td>
<td>QOL-D, QOL-AD NH</td>
<td>United Kingdom</td>
<td>421</td>
<td>79</td>
<td>85 (-)</td>
</tr>
<tr>
<td></td>
<td>DCM</td>
<td>United Kingdom</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Sloane et al. (2007)</td>
<td>DCM</td>
<td>United Kingdom</td>
<td>68</td>
<td>66.7</td>
<td>86.08 (7.04)</td>
</tr>
<tr>
<td>Villar et al. (2015)</td>
<td>DCM</td>
<td>Spain</td>
<td>42</td>
<td>67</td>
<td>80.6 (-)</td>
</tr>
<tr>
<td>Weiner et al. (1999)</td>
<td>QUALID</td>
<td>United States</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
</tbody>
</table>

*English language instrument used.

ADRQL, Alzheimer Disease Related Quality of Life; ADRQL revised, Alzheimer Disease Related Quality of Life revised version; DCM, Dementia Care Mapping; DQoL, Dementia quality of life; QOL-AD, Quality of Life in Alzheimer’s Disease; QOL-AD NH, Quality of Life in Alzheimer’s Disease nursing home version; QOL-D, Quality of Life in Dementia; QUALID, The quality of life in late-stage dementia; QUALIDEM, Quality of life instrument for proxy completion; MMSE, Mini Mental State Examination (Folstein et al., 1975); FAST, Functional Assessment Staging (Reisberg, 1988); GDS, Global Deterioration Scale (Reisberg et al., 1982); CDR, Clinical Dementia Rating Scale (Morris, 1997); – = Not reported.
<table>
<thead>
<tr>
<th>QOL INSTRUMENT</th>
<th>AUTHOR(S)</th>
<th>INTERNAL CONSISTENCY</th>
<th>RELIABILITY</th>
<th>MEASUREMENT ERROR</th>
<th>CONTENT VALIDITY</th>
<th>STRUCTURAL VALIDITY</th>
<th>HYPOTHESIS TESTING</th>
<th>CROSS-CULTURAL VALIDITY</th>
<th>CRITERION VALIDITY</th>
<th>RESPONSIVENESS</th>
</tr>
</thead>
<tbody>
<tr>
<td>DQoL</td>
<td>Edelman et al. (2005)</td>
<td>Good</td>
<td>•</td>
<td>•</td>
<td>Poor</td>
<td>Poor</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
<tr>
<td></td>
<td>Adler and Resnick (2010)</td>
<td>Poor</td>
<td>Fair</td>
<td>•</td>
<td>•</td>
<td>Fair</td>
<td>•</td>
<td>•</td>
<td>Poor</td>
<td>•</td>
</tr>
<tr>
<td></td>
<td>Moyle et al. (2012)</td>
<td>Poor</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>Fair</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
<tr>
<td></td>
<td>Sloane et al. (2005)</td>
<td>Poor</td>
<td>Fair</td>
<td>•</td>
<td>•</td>
<td>Poor</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
<tr>
<td>QOL-AD NH</td>
<td>Edelman et al. (2005)</td>
<td>Good</td>
<td>•</td>
<td>•</td>
<td>Poor</td>
<td>Poor</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
<tr>
<td></td>
<td>Moyle et al. (2012)</td>
<td>Poor</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>Fair</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
<tr>
<td></td>
<td>Sloane et al. (2005)</td>
<td>Poor</td>
<td>Fair</td>
<td>•</td>
<td>•</td>
<td>Poor</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
<tr>
<td></td>
<td>Crespo et al. (2011)</td>
<td>Poor</td>
<td>Fair</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>Poor</td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
<tr>
<td></td>
<td>Dichter et al. (2016b)</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>Fair</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
<tr>
<td></td>
<td>Beer et al. (2009)</td>
<td>Poor</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
<tr>
<td>QUALIDEM</td>
<td>Dichter et al. (2013)</td>
<td>Excellent</td>
<td>•</td>
<td>•</td>
<td>Excellent</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Bouman et al. (2011)</td>
<td>Excellent</td>
<td>Fair</td>
<td>•</td>
<td>•</td>
<td>Excellent</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
<tr>
<td></td>
<td>Dichter et al. (2011)</td>
<td>Good</td>
<td>•</td>
<td>•</td>
<td>Fair</td>
<td>Poor</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
<tr>
<td></td>
<td>Ettema et al. (2007b)</td>
<td>Poor</td>
<td>Good</td>
<td>•</td>
<td>Fair</td>
<td>Good</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
<tr>
<td></td>
<td>Dichter et al. (2016a)</td>
<td>Poor</td>
<td>Fair</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
<tr>
<td></td>
<td>Dichter et al. (2014)</td>
<td>•</td>
<td>Fair</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
<tr>
<td></td>
<td>Ettema et al. (2007a)</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>Fair</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
<tr>
<td>QOL INSTRUMENT</td>
<td>AUTHOR(S)</td>
<td>INTERNAL CONSISTENCY</td>
<td>RELIABILITY</td>
<td>MEASUREMENT ERROR</td>
<td>CONTENT VALIDITY</td>
<td>STRUCTURAL VALIDITY</td>
<td>HYPOTHESIS TESTING</td>
<td>CROSS-CULTURAL VALIDITY</td>
<td>CRITERION VALIDITY</td>
<td>RESPONSIVENESS</td>
</tr>
<tr>
<td>----------------</td>
<td>----------------------------------</td>
<td>----------------------</td>
<td>-------------</td>
<td>-------------------</td>
<td>------------------</td>
<td>---------------------</td>
<td>--------------------</td>
<td>--------------------------</td>
<td>---------------------</td>
<td>---------------</td>
</tr>
<tr>
<td>QUALID</td>
<td>Garre-Olmo et al. (2010)</td>
<td>Good</td>
<td>Fair</td>
<td>Fair</td>
<td>•</td>
<td>Fair</td>
<td>Fair</td>
<td>Poor</td>
<td>Poor</td>
<td>•</td>
</tr>
<tr>
<td></td>
<td>Mjorud et al. (2014)</td>
<td>Good</td>
<td>•</td>
<td>•</td>
<td>Excellent</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
<tr>
<td></td>
<td>Clare et al. (2014)</td>
<td>Fair</td>
<td>•</td>
<td>•</td>
<td>Fair</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
<tr>
<td></td>
<td>Falk et al. (2007)</td>
<td>Fair</td>
<td>Poor</td>
<td>•</td>
<td>Fair</td>
<td>Poor</td>
<td>Poor</td>
<td>Poor</td>
<td>Poor</td>
<td>Poor</td>
</tr>
<tr>
<td></td>
<td>Røen et al. (2015)</td>
<td>Poor</td>
<td>Fair</td>
<td>•</td>
<td>•</td>
<td>Poor</td>
<td>Poor</td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
<tr>
<td></td>
<td>Weiner et al. (1999)</td>
<td>Poor</td>
<td>Poor</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
<tr>
<td>ADRQL</td>
<td>Edelman et al. (2005)</td>
<td>Poor</td>
<td>•</td>
<td>•</td>
<td>Poor</td>
<td>Poor</td>
<td>Poor</td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
<tr>
<td></td>
<td>Sloane et al. (2005)</td>
<td>Poor</td>
<td>Fair</td>
<td>•</td>
<td>•</td>
<td>Poor</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
<tr>
<td></td>
<td>Kasper et al. (2009)</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
<tr>
<td>ADRQL revised</td>
<td>Kasper et al. (2009)</td>
<td>Poor</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
<tr>
<td>DCM</td>
<td>Villar et al. (2015)</td>
<td>Good</td>
<td>Poor</td>
<td>•</td>
<td>•</td>
<td>Fair</td>
<td>Fair</td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
<tr>
<td></td>
<td>Fossey et al. (2002)</td>
<td>Poor</td>
<td>Fair</td>
<td>•</td>
<td>•</td>
<td>Poor</td>
<td>Poor</td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
<tr>
<td></td>
<td>Sloane et al. (2007)</td>
<td>•</td>
<td>Fair</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
<tr>
<td></td>
<td>Sloane et al. (2005)</td>
<td>•</td>
<td>Fair</td>
<td>•</td>
<td>•</td>
<td>Poor</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
<tr>
<td></td>
<td>Edelman et al. (2005)</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>Poor</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
<tr>
<td>QOL-AD</td>
<td>León-Salas et al. (2011)</td>
<td>Poor</td>
<td>Poor</td>
<td>Fair</td>
<td>•</td>
<td>•</td>
<td>Fair</td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
<tr>
<td>QOL-D</td>
<td>Sloane et al. (2005)</td>
<td>Poor</td>
<td>Fair</td>
<td>•</td>
<td>•</td>
<td>Poor</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
</tbody>
</table>
reliability scores from a larger number of studies (0.69–0.95).

**Measurement error**
QUALID and QOL-AD were the only instruments to have measurement error assessed; both were given fair quality ratings.

**Content validity**
Content validity was assessed in QUALIDEM only, and was rated as fair quality. Structural validity was assessed in six instruments; the instruments most assessed were QUALID, rated in five studies with poor to excellent ratings, and QUALIDEM, rated in four studies with fair to excellent ratings. DQoL and DCM were rated in two studies with fair to poor ratings for both. QOL-AD NH and ADRQL were assessed once with poor ratings.

**Hypothesis testing**
Hypothesis testing was carried out for eight instruments. The most assessed instruments were QUALID and DCM, which were assessed in four studies with poor to fair ratings. DQoL and QOL-AD NH were assessed in three studies with poor to fair ratings each. ADRQL was assessed in two studies with poor ratings, QUALIDEM and QOL-AD were assessed once with fair ratings, QOL-D had one poor rating.

**Cross-cultural validity**
Cross-cultural validity was assessed for three instruments (QUALID, QOL-AD NH, and QUALIDEM). QUALID and QOL-AD NH were both assessed twice. QUALID had poor ratings, QOL-AD NH had a poor and a fair rating. QUALIDEM was assessed once and had a poor rating.

**Criterion validity**
Criterion validity was assessed for two instruments. QUALID was assessed twice, both with poor ratings, and DQoL was assessed once with a poor rating. Only the QUALID was assessed for responsiveness and this had a poor rating.

**Feasibility properties**
Feasibility properties were extracted for all instruments, except for the case of the QoL-D, in which the original questionnaire and associated materials could not be identified even after contacting the original author. Full details of availability and feasibility properties are presented in Table 3.

**Dedicated website**
Four of the instruments (QOL-AD, QUALIDEM, ADRQL revised, and DQoL) had dedicated websites where they could be accessed. The ADRQL revised and QOL-AD were both available via the online MAPI trust repository. All other instruments required contacting the original author to access.

**User guide**
A user guide was accessible for six of the instruments (QOL-AD, QOL-AD NH, QUALIDEM, DCM, ADEQL, ADRQL revised), with many also including additional instructions on the actual questionnaire. We were unable to identify any user instructions for the DQoL and QUALID.

**Cost**
All instruments were free to access for non-funded research. Additional charges applied for funded research and commercial users in the case of the ADRQL, ADQRL revised, QOL-AD and QOL-AD NH. A single instrument (DCM) was free to use, but required users to attend training, which had associated costs.

**Training**
Four instruments (DQoL, QOL-AD, QOL-AD NH, and QUALIDEM) did not require any formal training before use of the instrument, whilst the ADRQL and ADRQL revised recommended users to watch a free training video. Only the DCM instrument required users to attend a three-day training course. It was unclear whether the QUALID required training.

**Time to complete**
Time to complete the instrument ranged from 5 minutes (QUALIDEM, QOL-AD NH) to up to 6 hours (DCM).

**Time period to assess**
Three instruments captured “present” QoL (QOL-AD, QOL-AD NH, DCM), whilst the remaining instruments assessed QoL over the previous 1 or 2 weeks. The only exception is the DQoL which does not state the time period in which participants should be assessed.

**Specialist software**
No instrument appeared to require the use of specialist software for the administration, analysis, or interpretation of the instruments.
## Table 3. Availability and feasibility properties of identified instruments

<table>
<thead>
<tr>
<th>SCALE</th>
<th>AVAILABILITY</th>
<th>USABILITY</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>DEDICATED WEBSITE</td>
<td>USER GUIDE</td>
</tr>
<tr>
<td>DQoL</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>QOL-AD</td>
<td>No*</td>
<td>Yes, and instructions on instrument</td>
</tr>
<tr>
<td>QOL-AD NH</td>
<td>No</td>
<td>Yes, and instructions on instrument</td>
</tr>
<tr>
<td>QUALIDEM</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>QUALID</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>DCM</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>ADRQL</td>
<td>Yes</td>
<td>Yes, and instructions on instrument</td>
</tr>
<tr>
<td>ADRQL Revised</td>
<td>Yes*</td>
<td>Yes, and instructions on instrument</td>
</tr>
<tr>
<td>QoL-D</td>
<td>No</td>
<td>Not known</td>
</tr>
</tbody>
</table>

All information correct as of January 2018.

*Not a dedicated website but available from MAPI trust repository.
*a Non-funded academic researchers - free; Funded academic researchers - €300 per study and €50 per language; Commercial users – Royalty fees €1000 per study and €50 per language. Distribution fees €1000 per study and €50 per language.
*b Non-funded academic researchers — free; Funded academic researchers — €300 per study and €50 per language; Commercial users – Royalty fees 10,000 USD per study and 500 USD per language. Distribution fees €700 per study and €300 per language.
Scoring guide

All instruments had accessible scoring instructions.

Discussion

This study aimed to add usefully to existing reviews of the literature (Bowling et al., 2015; Ready and Ott, 2003). Aspden et al. (2014) by assessing the usability properties of instruments in care homes in addition to psychometric properties and study quality. This provides broader information on the pragmatic use of dementia-specific QoL instruments in care homes, which is likely to be of value to researchers, clinicians, and practitioners when selecting instruments for use both in research and in normal practice.

Twenty-five studies were identified that assessed nine dementia specific QoL instruments for use in care homes. This review highlights that even though QoL instruments exist for use in care homes, they may not be accessible or feasible for use. There was limited information about the psychometric properties of most instruments; as a consequence many elements were not assessed. COSMIN scores that were assessed had relatively low ratings with the majority of ratings being poor or fair. The instruments with the most psychometric evidence were QUALID and QUALIDEM. The QUALID instrument had the most extensive evaluation of the two, with seven domains assessed, however, QUALIDEM had better ratings for most of the assessed properties with more excellent and good ratings compared to QUALID. The QUALIDEM has previously been suggested as the best QoL instrument to use for people with dementia in care homes due to the comprehensive assessment of its measurement properties (Aspden et al., 2014).

The QOL-AD instrument is one of the most widely used QoL instruments due to its good psychometric properties and apparent ease of use (Bowling et al., 2015). However, findings from the review show limited use of the QOL-AD in care settings, with the original QOL-AD only identified in one study, and the nursing home version identified in six studies. The QOL-AD NH instrument, which is adapted from QOL-AD was assessed in six studies; however, the ratings were mainly poor or fair. These findings are in line with a previous review which found that the QOL-AD NH properties were poor because of small sample sizes and methodological quality of the studies (Aspden et al., 2014).

It is widely accepted that measuring QoL in care settings is important. This is usually carried out as a part of research with the focus being to understand changes to QoL over time or to assess outcomes of interventions (Clare et al., 2014b; Hoe et al., 2009). There is, however, an additional need and potential benefit in understanding the QoL of care home residents outside of research, in routine care practice. The lack of routine measurement of QoL in care homes will be in part due to differences in the current instruments and a lack of consensus about what a QoL instrument should contain. The frameworks of QoL in dementia underpinning each instrument influence and shape the instrument and its content (Missotten et al., 2016). The instruments discussed here are based on a small number of models. Most are based solely, or in part, on the work of Lawton (Missotten et al., 2016), which states that QoL in dementia is multidimensional and consists of both objective and subjective components. A health-related QoL definition, which only includes aspects of QoL affected by a health condition, and the adaptation-coping model, which is concerned with adaptation to the consequences of dementia (Dröes et al., 2010) have also been used. However, it is often not clear what conceptual framework instruments are based on, and what assumptions are made (Bowling et al., 2015; Missotten et al., 2016). This may dissuade some from using QoL instruments, and the differences in the content of instruments may influence whether or not specific instruments are used. This, coupled with what it is that practitioners, researchers, or clinicians want to gain from measuring QoL will influence the decision about what instrument to use. However a fundamental problem is likely to be limited access to and experience of QoL instruments as well as a lack of requirement by regulators to use such approaches. Most are difficult to find and access online; often requiring contacting the original author or registering with online instrument repositories to request access to specific instruments. This will be difficult for care staff who may not be aware of the different instruments available and may not have the skills and resources needed to identify them.

Routine use of QoL instruments might be of value at an individual resident level and also at an aggregate level providing insights into the QoL of residents as a whole and indirectly the home as a whole. QoL measurement has the potential to improve the provision of person-centered or holistic care by encouraging care staff to focus more on the individual and less on impairments of function (Edvardsson et al., 2014). The inclusion of an instrument to specifically measure QoL in normal care practice, and used in conjunction with existing care practices, procedures, and documentation, could enhance the opportunity to improve care quality by providing more person-centered information on residents and providing feedback on the effect of changes over time. In the US, a comprehensive assessment of a variety of care...
Second, dual measurement is dependent on an
measurement for that individual resident over time.
instrument, thus meaning a lack of consistent
overall, and changes to resident status (from no
dementia are in the majority, it may be prefera-
be used for people with dementia and another for
colleagues suggested that one instrument should
in an English-speaking sample and setting.
be determined from this review. Therefore,
dementia-specific QoL instruments may be a better fit for care home resi-
dents, even those without dementia, than generic
This can be posed as an empirical
question and requires research to be done to test
this hypothesis. On balance we believe a strategy of
dual measurement may introduce more measure-
ment error than it prevents with the addition of
unnecessary complexity.

Despite differences in the content of instruments
and the findings presented in this review, the use of
existing QoL instruments is preferable to using no
QoL instruments at all. They provide potentially
valuable information compared with not using
them. There is, however, a need for the development
of instruments that are developed and/or validated
to measure QoL in care home settings and that
are accessible, easy to use, and contain appropriate
questions. The future development and adaptation
of instruments needs to consider these care-home
specific points to create instruments that are of
demonstrable benefit and lead to the wider use of
QoL instruments in care homes.

In this review we do not recommend the use of
one instrument as being the best or most appropriate
to use. Instead, because there is no gold standard
instrument, we argue that there are factors other
than psychometric properties that should be consid-
ered when deciding on an appropriate instrument to
use in care settings, particularly if the care staff will
be involved in collecting data. We conclude that
new instruments need to be developed either from
existing measures or de novo that are specific to the
care home setting, that have acceptable psychomet-
ric properties and that are readily accessible and
usable in these settings. Deciding on what instru-
ment is the most useful in care settings is difficult
given the differences in care homes themselves.

features is completed for residents in long-term care
facilities. The Resident Assessment Instrument
(RAI-NH) is completed periodically to collect and
record relevant information about resident care. The
aim of its use is to inform holistic care and ensure
good QoL and quality of care. Implementation and
use of the RAI-NH has been shown to be associated
with improvements in resident outcomes (Fries et al.,
1997; Phillips et al., 1997). The integration of a QoL
instrument into such a system and used regularly
could enhance the assessment of resident outcomes
by providing complementary information and there-
fore a broader understanding of wellbeing of the
resident and when aggregated further insights into
the quality of the care home. In the UK and many
other countries there is no such mandated combined
assessment schedule in the care sector and no culture
of routine measurement; resident and care factors
are documented by staff, in a less systematic manner
than in the RAI-NH. Therefore, integrating QoL
measurement into routine care practice would need a
different approach. Regardless of the systems that
currently exist in homes, such action would require
appropriate instruments that are accessible for and
usable by care home staff. Of all the instruments that
were available, the QOL-AD, QUALIDEM, and
ADRQL revised were accessible via simple internet
searches and had a dedicated web page. Importantly,
many of the measures were difficult to access, and
required contacting the original author (e.g., DQoL).
Overall, the QUALIDEM was assessed as the most
accessible; being free to use, with all information
regarding administration and scoring available in
an extensive user guide. However, although an
English language version of QUALIDEM is available,
it has not been validated in an English sample as far as
could be determined from this review. Therefore,
despite QUALIDEM being the most psychometri-
cally robust and accessible instrument there is still
some concern regarding whether it is suitable for use
in an English-speaking sample and setting.

In this review we focus on dementia-specific QoL
instruments. A previous review by Aspden and
colleagues suggested that one instrument should
be used for people with dementia and another for
people without dementia (Aspden et al., 2014). We
understand the clarity of purpose behind this, but
believe that in routine care practice, when those
with dementia are in the majority, it may be prefera-
ble to use a single dementia-specific instrument for
three main reasons. First, using two different instru-
ments makes it difficult to aggregate scores for homes
overall, and changes to resident status (from no
diagnosis to diagnosis) would require a change in
instrument, thus meaning a lack of consistent
measurement for that individual resident over time.
Second, dual measurement is dependent on an
accurate diagnosis of a dementia having been
made for each resident and this is unlikely to be the
case. Diagnosis rates in the UK are amongst the
highest in the world but are still only around 67%
(Department of Health, 2016). Many without a
formal diagnosis of dementia in care homes will in
fact have dementia, and evidence of high prevalence
of dementia in care homes mean that most residents
would be appropriate for the use of a dementia-
specific QoL instrument (Jagger and Lindsay, 1997;
Livingston et al., 2017). Finally, the questions con-
tained in QoL instruments need to be appropriate for
people with dementia and for use in care settings.
Few generic measures of QoL have been developed
specifically for use in care home settings, and very
few have been validated in this setting. They can,
therefore, contain inappropriate questions, often
reflecting the opportunity to perform a function
rather than the ability to perform it (Hall et al., 2011).
We believe that the content of dementia-specific QoL
instruments may be a better fit for care home resi-
dents, even those without dementia, than generic
instruments. This can be posed as an empirical
question and requires research to be done to test
this hypothesis. On balance we believe a strategy of
dual measurement may introduce more measure-
ment error than it prevents with the addition of
unnecessary complexity.
Issues of time and resources and the benefit that accrues from their use will always be to the fore, so data on these parameters are needed. The development or adaptation of instruments that can be used in care homes by both researchers and by care staff routinely in normal care practice would be useful. If adapted, the content of such instruments would need to remain largely the same to maintain reliability and validity of the instrument. However, the format or layout of the instrument questions, and user guides or instructions would need to be adapted for use in care homes and by care staff. Researchers will be specifically trained and experienced in using standardised instruments and are likely to have more time to complete instruments. Care staff are likely to need more specific guidelines to follow and a layout that reduces the risk of recording error and missing data.

It is important to highlight several limitations of this review. First, this review did not include all measures of dementia-specific QoL, because they did not meet inclusion criteria of this review. For example, the Philadelphia Geriatric Centre Affect Rating Scale (PGC-ARS) (Lawton et al., 1996) and the Resident and Staff Observation Checklist Quality of Life measure (RSOC-QOL) (Sloane et al., 1991) were not included in this review despite being included in a review by Aspden and colleagues as the instruments were not assessed to be measures of QoL (Aspden et al., 2014). Second, there is no established means of assessing the usability of instruments in care homes, and therefore interpretation of what is considered “accessible” or “practical” may differ depending on the user. Third, this review focused on disease-specific instruments of QoL and makes the assumption that the majority of residents in a care home would have dementia; this may not be the case and could affect the validity of instruments. Finally, this review only included articles published in English, meaning that non-English instruments may not have been captured.

Conclusion

The number of high quality easily accessible instruments available for use in care homes is low. In general, the psychometric analysis of instruments revealed them to function poorly in care home settings with a limited number of psychometric elements assessed for each instrument in any single study. Furthermore, the quality ratings of each instrument were weak with very few receiving excellent or good ratings. Instruments with the best psychometric assessments were not necessarily easily accessible, and those that were readily available had some of the poorest quality and most limited assessment. The findings of this review indicate that there is a need for further large and well-designed evaluations of the properties of these instruments when used in care homes, including head-to-head comparisons. Further work is also needed in order to ensure that any instruments are made easily available for use by care staff if they are to be used in a non-research or clinical capacity. Due to poor psychometric and feasibility properties, none of the instruments currently available seem suited to routine use by care staff in care homes. Further development of existing instruments or the generation of new methods for the measurement of QoL in care homes would be of potential value given the need to assure and improve the quality of care of residents in care homes with dementia.

Conflict of interest

None.

Description of authors’ roles

LH completed the search, assessed the articles, and wrote the first draft of the review. NF and TEP assisted with article assessment and screening, and provided inputs and amendments to the review. NT and SB provided input and amendments throughout the development of the review.

Supplementary material

To view supplementary material for this article, please visit https://doi.org/10.1017/S1041610218002259.

References


Department of Health, Prime Minister’s Office. (2015). Prime Minister’s challenge on dementia. Available at: https://www.gov.uk/government/publications/prime-ministers-challenge-on-dementia-2020-


Quality of life instruments in care settings


Mokkink, L. B. et al. (2010). The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *Journal of Clinical Epidemiology*, 63, 737–745.


