Psychometric properties of carer-reported outcome measures in palliative care: A systematic review

Charlotte TJ Michels1,2, Mary Boulton3, Astrid Adams4, Bee Wee3,4,5 and Michele Peters1

Abstract

Background: Informal carers face many challenges in caring for patients with palliative care needs. Selecting suitable valid and reliable outcome measures to determine the impact of caring and carers’ outcomes is a common problem.

Aim: To identify outcome measures used for informal carers looking after patients with palliative care needs, and to evaluate the measures’ psychometric properties.

Design: A systematic review was conducted. The studies identified were evaluated by independent reviewers (C.T.J.M., M.B., M.P.). Data regarding study characteristics and psychometric properties of the measures were extracted and evaluated. Good psychometric properties indicate a high-quality measure.

Data sources: The search was conducted, unrestricted to publication year, in the following electronic databases: Applied Social Sciences Index and Abstracts, Cumulative Index to Nursing and Allied Health Literature, The Cochrane Library, EMBASE, PubMed, PsycINFO, Social Sciences Citation Index and Sociological Abstracts.

Results: Our systematic search revealed 4505 potential relevant studies, of which 112 studies met the inclusion criteria using 38 carer measures for informal carers of patients with palliative care needs. Psychometric properties were reported in only 46% (n = 52) of the studies, in relation to 24 measures. Where psychometric data were reported, the focus was mainly on internal consistency (n = 45, 87%), construct validity (n = 27, 52%) and/or reliability (n = 14, 27%). Of these, 24 measures, only four (17%) had been formally validated in informal carers in palliative care.

Conclusion: A broad range of outcome measures have been used for informal carers of patients with palliative care needs. Little formal psychometric testing has been undertaken. Furthermore, development and refinement of measures in this field is required.

Keywords

Palliative care, carers, outcome measures, psychometrics, systematic review

What is already known about the topic?

- The involvement of informal carers is essential for the provision of palliative care, but informal caregiving can have a major impact on carers’ outcomes.
- Studies of informal carer outcomes use a wide range of endpoints.
- Selecting suitable and appropriate carer outcome measures seems problematic.
What this paper adds?

- An increasing number of studies are conducted in informal carers looking after patients with palliative care needs.
- Only four outcome measures have been formally developed and validated within this population, and limited psychometric information is available on most measures.
- While there has been an increasing trend since 2008 in the use of outcome measures for informal carers in palliative care research, most measures used in these studies were developed more than 20 years earlier and may not adhere to current standards for measure development.

Implications for practice, theory or policy

- Existing carer outcome measures need to be validated for the palliative care setting and new measures need to be developed in accordance with current guidelines in order to meet the requirements of the growing number of studies, including intervention studies, of informal carers looking after patients with palliative care needs.
- When using an existing outcome measure, the authors should report their rationale for selecting it and should refer to the publications that report the original development of the measure.
- Interventions for supporting informal carers should be evaluated using outcome measures for which appropriate psychometric properties have been reported before they are implemented as policy.

Introduction

The World Health Organization (WHO)\(^1\) defines palliative care as an approach that focuses on the quality of life of patients and their relatives facing problems associated with life-threatening illness, through prevention and relief of suffering. Annually, around 20 million people worldwide need palliative care,\(^2\) and an ageing population and increases in long-term conditions mean that need is likely to continue to rise.\(^3,4\)

Informal carers make an important contribution in the provision of palliative care and are regarded as integral to its delivery.\(^5,6\) Informal carers are defined as carers who are not financially compensated for their services typically spouses, children, siblings or friends.\(^7\) In 2011, the contribution of approximately six million informal carers in the United Kingdom was estimated at the equivalent of £119 billion a year.\(^8\) About half a million people are caring for patients during the end-of-life phase and this number is expected to increase to 3.4 million in the coming 30 years.\(^9\)

Palliative care has become an important component of health care, and policy makers are putting more emphasis on informal carers.\(^10\) Informal caregiving may provide emotional benefits and togetherness for carers,\(^11,12\) but it also involves considerable challenges including adverse psychological, physical, social and financial consequences.\(^13,14\) Studies indicate that informal caregiving affects carers’ wellbeing and their own health resulting in isolation, fatigue, sleeping problems, exhaustion, weight loss, depression and anxiety.\(^15–19\) It is therefore important that carer outcomes are assessed in order to be able to provide effective support and to reduce negative consequences of caregiving. Carer outcomes refer to a range of concepts including quality of life, burden and strain. While these terms are not well defined and frequently get used interchangeably, it is generally accepted that they comprise multiple dimensions such as physical impact, mental strain and social functioning. It has been argued that quality of life is a broader concept as it assesses a wider spectrum of wellbeing, whereas burden and strain suggest a more direct measure of duty of care.\(^20\)

Evidence on effective strategies to reduce the burden of caring and improve their quality of life of informal carers is limited.\(^21,22\) Although interventions have been developed that aim to improve outcomes for informal carers, their results are difficult to compare as studies focus on a wide range of endpoints.\(^23\) One systematic review identified 62 questionnaires used among informal family carers in various palliative care settings.\(^24\) These questionnaires included instruments on carer satisfaction, experience (of health services and support), needs bereavement and outcomes. Previous reviews on interventions for informal carers concluded that it was unclear what kind of support was beneficial, partly due to the lack of appropriate outcome measures.\(^21,25\)

In order to assess the impact of the caring role on carers, an appropriate choice of outcome measures is required; however, selecting suitable and appropriate measures seems a common problem.\(^25–27\) This requires reliable and valid measures with robust psychometric properties, which are appropriate for a palliative care context, as this forms the foundation for evaluating caregiver interventions.

This systematic review aimed to identify and evaluate outcome measures that have been used for informal carers in palliative care studies. The measures used in palliative care are described and their psychometric properties (e.g. reliability, validity, feasibility and precision), when available, are evaluated.
Study selection

All types of multidimensional measures (generic, carer-specific for any condition and carer-specific for patients with a specific condition) were eligible for inclusion. The study focused on multidimensional measures as we were interested in measures that assess the overall impact of caring in palliative care rather than measures that assess one specific dimension of outcome or impact. A study was included if all of the following were fulfilled: (1) the study used a self-reported multidimensional measure that assessed caregiver outcomes (i.e. burden, strain or quality of life), (2) measures were directed at unpaid informal carers (e.g. spouse, relatives, siblings, friends or neighbours), (3) the patients they supported were diagnosed with an advanced progressive illness or were receiving palliative care (end-of-life care, terminal care or hospice care), (4) both carers and patients were ≥18 years old and (5) the study was reported in English.

A study was excluded if any of the following were fulfilled: (1) only unidimensional measures were used; (2) only subscales or individual items and not the full measure were included; (3) only clinician-assessed measures or patient-reported measures were used; (4) all measures completed by carers were on behalf of the patient or (5) it was a qualitative study, comment, editorial, protocol, conference article or grey literature. There were no restrictions regarding publication date and research methods.
Data extraction and analysis

After retrieving all records, the duplicates were removed. All studies were initially screened on the basis of title and abstract, and then on the basis of full-text. Three authors (C.T.J.M., M.B. and M.P.) independently assessed the eligibility of studies: C.T.J.M. assessed all articles, M.B. and M.P. each assessed half of the articles. Any uncertainties were discussed with the other two authors (A.A. and B.W.) and resolved by consensus. C.T.J.M. extracted the data on study characteristics (publication year, country, sample size, research setting, type of disease, intended outcome measure and information on measure) and psychometric characteristics. The following information on psychometrics was collected: content validity, internal consistency, construct validity, reproducibility (agreement and reliability), responsiveness, floor or ceiling effects, acceptability and feasibility. As guidance, we used the definitions given by Terwee et al.²⁹ and Fitzpatrick et al.³⁰ Additionally, when an included study did not report any psychometric information but referred to other articles regarding a measure or its psychometric values, we assessed these additional articles in order to evaluate the evidence they provided.

Results

Our electronic search, performed on 4 September 2014, identified 8569 studies. Figure 1 provides an overview of the number of studies identified at each stage of the search. After duplicates were removed, 4505 studies were screened on the basis of titles and abstracts, and 231 studies were screened on the basis of full text. This identified 112 studies using 38 different measures for informal carers in palliative care.

Study and measure characteristics

A total of 112 studies (18 randomized controlled trials (RCTs), 78 observational studies and 16 methodological studies) were included. The methodological studies included translation, development and validation studies about an outcome measure for informal carers in palliative care.
care. The patient population mainly consisted of cancer patients (n=67, 60%) or a mixture of conditions (n=29, 26%). Of the studies, 37% were conducted in the United States. Most studies included a mix of spouses, children, parents or friends (n=99, 88%) and a small number of studies included only spouse carers (n=4, 3%).

Most studies used only one outcome measure that fit our selection criteria (n=91, 81%) and 19% of the studies administered two outcome measures to carers. Studies mainly used carer-specific measures only (n=69, 62%), a quarter used a generic measure (n=29, 26%), and 14 studies used both types (i.e. generic and carer-specific). In total, 38 measures were identified, including 25 carer-specific measures and 13 generic measures. The main study characteristics are presented in Table 2 and in detail in Supplement 1.

The most frequently used generic measure was the SF-36 (n=16, 14%). The most frequently used carer-specific measures were the Caregiver Reaction Assessment (n=21, 19%), Caregiver Quality of Life Index–Cancer (n=14, 13%) and the Zarit Burden Inventory (n=10, 9%). The primary focus of studies using a carer-specific measure was burden (n=14, 13%), followed by quality of life (n=8, 7%) and strain (n=3, 2.6%). An overview of the identified measures and their frequency of use are presented in Table 3.

**Psychometrics of measures**

More than half of the 112 (n=60, 54%) studies reported no information on psychometric properties. The 52 (46%) studies that did included 33 observational studies, 15 methodological studies and 4 RCTs. Psychometric data were available for only 23 of the 38 measures including 7 generic measures (i.e. McGill Quality of Life Questionnaire, World Health Organization Quality of Life, Quality of Life Scale, Quality of Life Index, SF-36, SF-12 and Swedish Health-Related Quality of life) and 17 carer-specific measures. These measures consisted of 4–64 items, with a median of 16 items. Table 4 presents an overview of the 24 measures with the available psychometric information. This consisted mainly of information on the Cronbach’s alpha (n=45, 40%), construct validity (n=27, 24%), reliability (n=14, 12%), content validity (n=8, 7%), responsiveness (n=8, 7%) and acceptability and feasibility (n=8, 7%).

Of the 24 measures, four were originally developed in a palliative care context, that is, the Quality of Life in Life-Threatening Illness–Family Carer Version (QOLLTI-F), the Family Appraisal of Caregiving Questionnaire for Palliative Care (FACQ-PC), the Caregiver Burden Scale in end-of-life-care (CBS-EOLC) and the Caregiver Quality of Life Index (CQOLI). The content validity (which examines the extent to which the concepts of interest are represented by the items) and internal consistency (which measures the extent to which items in a scale are inter correlated) construct validity (the extent to which scores relate to other similar measured concepts) were adequate in all four measures. Interpretability (the degree to which one can assign qualitative meaning to quantitative scores) was not reported in all four studies. The reliability (which concerns the degree to which repeated measurements in stable persons provide similar answers) was positive in two measures and negative for FACQ-PC. Floor and ceiling effects (considered to be present if more than 15% of respondents achieved the lowest or highest possible score, indicating that it is likely that extreme items are missing in the lower or upper ends of the scale) was negative for QOLLTI-F and not reported for the other three measures.

For studies (n=60) that did not report psychometric properties but referred to previous publications about the measure, C.T.J.M. additionally extracted psychometric information from the referenced articles (see Supplement 2). An additional 139 references were assessed for study type, study population and psychometric properties. Although this provided information on how the measures were originally developed, it did not result in additional psychometric information for the measures in the context of carers in a palliative care setting.

**Discussion**

The aim of this systematic review was to identify and evaluate the psychometric properties of self-reported measures used in informal carers in palliative care studies. A total of 112 studies were found, which used 38 different outcome measures for informal carers. The most commonly used generic measure was the SF-36 and the most commonly used carer-specific measure was the Caregiver Reaction Assessment. Psychometric information was available for only 24 carer outcome measures (52 studies). We identified only four measures that were formally tested in a palliative care context.

Measures were mainly used in descriptive studies (n=78) and the overall study sample sizes tended to be quite small. This could be due to methodological and structural challenges in palliative care research. For example, uncertainties in patients’ prognosis, heterogeneity of the palliative care population, relatively small palliative care centres, ethical concerns or attrition of patients during the study could inhibit research in palliative care.

We noted an increasing trend in the use of measures in informal carers in palliative care. The majority of the included studies were published relatively recently, with more than 70% published since 2008. However, the majority of measures were developed much longer ago, including the most frequently used such as the Caregiver Reaction Assessment and the Zarit Burden Interview. It is therefore unclear whether measures adhere to the current...
Table 2. Study characteristics of the included studies (n = 112).

<table>
<thead>
<tr>
<th>Study characteristic</th>
<th>Number of studies (%)</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Type of study</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Methodological</td>
<td>16 (15%)</td>
<td>31–46</td>
</tr>
<tr>
<td>Observational</td>
<td>78 (70%)</td>
<td>17,47–123</td>
</tr>
<tr>
<td>RCT</td>
<td>18 (16%)</td>
<td>124–141</td>
</tr>
<tr>
<td><strong>Country</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Australia</td>
<td>12 (11%)</td>
<td>17,35,45,50,56,57,82,83,89,123,133,138</td>
</tr>
<tr>
<td>Canada</td>
<td>11 (10%)</td>
<td>34,36,39,52,53,55,62,63,77,78,139</td>
</tr>
<tr>
<td>Norway</td>
<td>8 (7%)</td>
<td>72–76,103,118,119</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>7 (6%)</td>
<td>37,64,80,100,126,137,140</td>
</tr>
<tr>
<td>Other (e.g. Brazil, China, Germany, Spain, The Netherlands, Japan, Korea, Sweden and Taiwan)</td>
<td>37 (33%)</td>
<td>31–33,38,40–44,48,49,54,58,67,69,71,79,81,85,91–93,97,99,101,104,109–116,122,127,130</td>
</tr>
<tr>
<td><strong>Study population</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mixture of informal carers (e.g. spouse, child and parent)</td>
<td>99 (88%)</td>
<td>17,31–37,39–47,49,51–63,65–74,76–81,83–92,94,96–120,122–125,127,129,130,132,134–141</td>
</tr>
<tr>
<td>Not reported</td>
<td>9 (8%)</td>
<td>38,64,75,82,95,12,126,128,133</td>
</tr>
<tr>
<td>Spouse</td>
<td>4 (3%)</td>
<td>48,50,93,131</td>
</tr>
<tr>
<td><strong>Patient population (disease)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other (e.g. ALS, ESRD, dementia, heart failure, MND and MS)</td>
<td>10 (9%)</td>
<td>33,49,50,64,79,93,107,121,126,137</td>
</tr>
<tr>
<td>Not reported</td>
<td>6 (5%)</td>
<td>36,60,61,117,125,141</td>
</tr>
<tr>
<td><strong>Sample size study population</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N &lt; 50</td>
<td>22 (20%)</td>
<td>47,48,50,51,55–57,59,60,63,65,79,90,95,96,100,107–109,121,125,137</td>
</tr>
<tr>
<td>N = 101–200</td>
<td>31 (28%)</td>
<td>33,35,39,42,46,58,61,64,66,69,70,76,78,82,85,89,98,99,101,104,111,115,118,119,124,127,130,133,135,139</td>
</tr>
<tr>
<td>Not reported</td>
<td>1 (1%)</td>
<td>126</td>
</tr>
<tr>
<td><strong>Type of measure</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Generic measure only</td>
<td>29 (26%)</td>
<td>37,40,42,55,63,64,67–69,72,73,76,89,90,93,97,101,103,107,109,110,119–121,127,130,132,137,141</td>
</tr>
<tr>
<td>Both generic and carer-specific measure</td>
<td>14 (12%)</td>
<td>33,49,65,71,77,99,102,108,118,123,128,135,136,139</td>
</tr>
<tr>
<td><strong>Number of measures asked in study</strong></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

RCT: randomized controlled trial; ALS: amyotrophic lateral sclerosis; ESRD: end-stage renal disease; MND: motor neurone disease; MS: multiple sclerosis.
development guidelines, such as those set by the Food and Drug Administration for patient-reported outcome measures. Evaluating publications on the development of these outcome measures was beyond the scope of our review, and the information would have been of limited value as the measures were mainly developed in other carer populations.

Due to the wide range of identified carer outcome measures and the variety of versions of the measures (e.g. Zarit Burden Interview; Table 3), it is difficult to draw overall conclusions about psychometric properties. The most commonly reported psychometric information was Cronbach’s alpha ($n = 45$, 40%), which is a psychometric property that is commonly used, relatively easy to calculate and easy to interpret. In all, 60 did not report any psychometric information. It was not expected that all studies would contain psychometric information, as the lack of psychometrics was not an exclusion criterion. For studies that did not report psychometric properties but referred to previous publications about the measure, we screened an additional 139 references for information on psychometrics. However, these resulted in limited extra psychometric data, and none of the studies met the inclusion criteria of this systematic review.

Although psychometric information was generally limited, it was even more limited in relation to some psychometric properties such as responsiveness. Responsiveness (or sensitivity to change) is particularly important to highlight as carer-reported outcome measures may be used to assess the effectiveness of interventions. Interventions to support carers in palliative care settings are likely to be complex and require measures that are able to detect change following the intervention.

We identified only four carer-specific measures that were formally developed and tested in this population: QOLLTI-F, FACQ-PC, CBS-EOLC and CQLI. Regarding the generic measures, none have been formally validated in this carer population but we found psychometric information on seven measures. As

<table>
<thead>
<tr>
<th>Measures</th>
<th>Number of studies</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Generic measures</td>
<td>SF-36</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>SF-12</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>SF-8*</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>EORTC QLQ-C30*</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>EQ-SD*</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>QOLS</td>
<td>3</td>
</tr>
<tr>
<td>Other (i.e. MS*, MQOL, SWED-QOL, OQOLI*, QOLI, WHOQOL and WHOQOL-BREF*)</td>
<td>9</td>
<td>42,93,101,107,109,127,128,130</td>
</tr>
<tr>
<td></td>
<td>ZBI (including 4 item, 6 item, 8 item, 12 item, 22 item 29 item version)</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>CBS</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>Other (i.e. BASC*, BCOS, BSFC, CBS-EOLC, CBI, CI, FAC5*, HP*, MBCBS, RCAS* and BIC)</td>
<td>15</td>
</tr>
<tr>
<td>Quality of life</td>
<td>CQOLI–Cancer</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>CQOLI–Revised</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>QOLLTI-F</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>Other (i.e. AQOL-EOL*, CH-QOL-F*, FACT, HQOLI* and QOL–Family*)</td>
<td>8</td>
</tr>
<tr>
<td>Strain</td>
<td>CSI</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>FACQ-PC</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>FSQ*</td>
<td>1</td>
</tr>
</tbody>
</table>


*No information reported and available on psychometric properties.
Table 4  Identified psychometric information in studies identified from the search (n=52).

<table>
<thead>
<tr>
<th>Measure</th>
<th>Study</th>
<th>Type of study</th>
<th>No. of items</th>
<th>Psychometric information</th>
<th>Measure references cited by study</th>
</tr>
</thead>
<tbody>
<tr>
<td>BCOS</td>
<td>Buscemi et al 34</td>
<td>OBS 15 items</td>
<td>α=0.75</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td></td>
<td>Govina et al38</td>
<td>MES 15 items</td>
<td>α=0.83</td>
<td>BCOS-LASA (r=0.7), BCOS-G-HADS (r=–0.52), Criterion validity r=0.57</td>
<td>ICC=0.985, ITC=0.47–0.76</td>
</tr>
<tr>
<td>BSFC</td>
<td>Brogaard et al 31</td>
<td>MES 28 items</td>
<td>α=0.91</td>
<td>Social isolation (r =0.33, p=0.01), ITC=0.02–0.72, Dyspnoea (r=0.32, p=0.01), Self-reported health (r=0.03, p=0.80)</td>
<td>Tendency towards floor effect.</td>
</tr>
<tr>
<td>BIC</td>
<td>Misawa et al11</td>
<td>MES 11 items</td>
<td>NR</td>
<td>“CRA subscales strongly correlated with supposed subscales of BIC”</td>
<td>NR</td>
</tr>
<tr>
<td>CBS</td>
<td>Akinci and Pinar79</td>
<td>MES 22 items</td>
<td>α=0.91</td>
<td>CFA=0.43–0.81 (All CBS factors correlated with each other in positive direction. All sub dimensions scores were negatively correlated with SF-36 (–0.58; –0.65)</td>
<td>ICC=0.985, ITC=0.37–0.70</td>
</tr>
<tr>
<td>CBI</td>
<td>Merluzzi et al 93</td>
<td>OBS 24 items</td>
<td>α=0.88</td>
<td>Correlations CBI-CGI factors 2, 3 and 4 (p&lt;0.05). Strongest correlations for both PSS-CBI were with factors 3+4 of the CGI.</td>
<td>NR</td>
</tr>
<tr>
<td>CBS-EOLC</td>
<td>Dumont et al36</td>
<td>MES 16 items</td>
<td>α=0.95</td>
<td>Construct validity= Most inter-item associations were consistent with conceptual framework qualitative study. Divergent validity=Interscale-correlations ZBI=0.72 (p&lt;0.01), POMS (fatigue)=0.69 (p&lt;0.01), POMS (vigour)=0.27(p&lt;0.05). Explaining overall variance = 64.8%</td>
<td>Sensitivity showed consistent associations with EOGO and unmet needs.</td>
</tr>
<tr>
<td>CIS</td>
<td>Cameron et al55</td>
<td>OBS 14 items</td>
<td>α=0.87</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>Measure</td>
<td>Study</td>
<td>Type of study</td>
<td>No. of items</td>
<td>Psychometric information</td>
<td>Reliability</td>
</tr>
<tr>
<td>---------</td>
<td>-------</td>
<td>---------------</td>
<td>--------------</td>
<td>--------------------------</td>
<td>-------------</td>
</tr>
<tr>
<td>CQOLI</td>
<td>McMillan and Mahon 1994</td>
<td>MES</td>
<td>4 items</td>
<td>Expert panel, α=0.76–0.88</td>
<td>No test–retest differences</td>
</tr>
<tr>
<td></td>
<td>Connell et al 1997</td>
<td>OBS</td>
<td>35 items</td>
<td>NR, α=0.90</td>
<td>Test–retest =0.95</td>
</tr>
<tr>
<td></td>
<td>Delgado-Guay et al 1999</td>
<td>OBS</td>
<td>35 items</td>
<td>NR, α=0.91</td>
<td>Test–retest =0.95</td>
</tr>
<tr>
<td></td>
<td>Meyers and Gray 1999</td>
<td>OBS</td>
<td>35 items</td>
<td>Expert panel (89%), α=0.87–0.90</td>
<td>Test–retest =0.95</td>
</tr>
<tr>
<td></td>
<td>Tang et al 1999</td>
<td>OBS</td>
<td>35 items</td>
<td>NR, α=0.91</td>
<td>Test–retest =0.95</td>
</tr>
<tr>
<td></td>
<td>Tang et al 1999 (Mandarin version)</td>
<td>MES</td>
<td>35 items</td>
<td>Translated, α=0.87</td>
<td>EFA showed 7 underlying factors explaining 48.15% of the variance. Caregivers’ QOL was inversely related to both patients’ (F=0.90, p=0.008) and caregivers’ pain (t=−4.22, p&lt;0.001). Correlation CQOLC-M–MOS-SS scores (r=0.26, P&lt;0.01), CQOLC-M–SWBS scores (r=0.30, P&lt;0.01).</td>
</tr>
<tr>
<td></td>
<td>Tang 2009</td>
<td>OBS</td>
<td>35 items</td>
<td>NR, α=0.90</td>
<td>Test–retest =0.95</td>
</tr>
<tr>
<td></td>
<td>Weitzner 1999</td>
<td>MES</td>
<td>35 items</td>
<td>Expert panel, α=0.91</td>
<td>Correlations with mental health (r=0.64), emotional distress (r=−0.52), burden (r=−0.65) and patient’s performance (r=−0.47), physical health (r=−0.13), social support (r=0.22) and social desirability (r=0.08), SF-36–CQOLC correlations social support and social desirability were low range (range=0.08–0.20) Correlated with SF-36, BDI, ECOG, STAI, CBS, MSPSS, and MCSDS.</td>
</tr>
</tbody>
</table>

(Continued)
Table 4. (Continued)

<table>
<thead>
<tr>
<th>Measure</th>
<th>Study</th>
<th>Study Type</th>
<th>No. of items</th>
<th>Psychometric information</th>
<th>Measure references cited by study</th>
<th>Original validation studies of measure (study population)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CQOLI-R</td>
<td>Wittenberg, Lyles et al</td>
<td>RCT</td>
<td>4 items</td>
<td>α=0.77</td>
<td>Reliability=0.94</td>
<td>Courtney 2005, McMillan 1994 (Caregivers of hospice cancer patients)</td>
</tr>
<tr>
<td>CRA</td>
<td>Andrews</td>
<td>OBS</td>
<td>24 items</td>
<td>α=0.57-0.85</td>
<td>Correlated with CES-D scale.</td>
<td>Given 1992, Given 1993, Radloff 1997 (Caregivers of persons with physical impairments or dementia)</td>
</tr>
<tr>
<td></td>
<td>Grov et al</td>
<td>OBS</td>
<td>24 items</td>
<td>α=0.76-0.83</td>
<td>The initial rotation indicated that a number of items (3, 5, 7, 10, 16, 18) had high loadings. Items with high loading on more than one component were removed. A third and final Principal Component Analysis (PCA) was performed on the remaining items indicating extraction of 5 components, which accounted for 23.9%, 16.7%, 10.1%, 7.9%, and 7.2% of the total variance.</td>
<td>Given 1992, Grov 2006, Kinsella 1998, Nijboer 1999</td>
</tr>
<tr>
<td>Misawa et al</td>
<td>MES (Japanese version)</td>
<td>MES</td>
<td>24 items</td>
<td>α=0.73-0.89</td>
<td>EFA= 5 factors were extracted, and the cumulative proportion of variance explained was 76.6%. Correlations Impact on schedule–time dependent burden (r=0.75), Caregiver’s self-esteem–emotional burden (r=0.66), Caregiver’s self-esteem–Existential burden (r=0.54), Impact on health–physical burden (r=0.75)</td>
<td>No ceiling effect or floor effect.</td>
</tr>
<tr>
<td>Morishita and Kamibeppu</td>
<td>OBS</td>
<td>24 items</td>
<td>α=0.83–0.91</td>
<td>NR</td>
<td>NR</td>
<td>Given 1992, Misawa 2009</td>
</tr>
<tr>
<td>Stein et al</td>
<td>RCT</td>
<td>NR</td>
<td>NR</td>
<td>α=0.82</td>
<td>NR</td>
<td>Given 1992</td>
</tr>
<tr>
<td>Tang</td>
<td>OBS</td>
<td>NR</td>
<td>NR</td>
<td>α=0.90</td>
<td>NR</td>
<td>Given 1992, Nijboer 1999</td>
</tr>
<tr>
<td>Tang et al</td>
<td>OBS</td>
<td>NR</td>
<td>NR</td>
<td>α=0.88</td>
<td>NR</td>
<td>Given 1992</td>
</tr>
</tbody>
</table>
Table 4. (Continued)

<table>
<thead>
<tr>
<th>Measure</th>
<th>Study</th>
<th>Type of study</th>
<th>No. of items</th>
<th>Psychometric information</th>
<th>Measure references cited by study</th>
<th>Original validation studies of measure (study population)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Content Validity</td>
<td>Internal consistency</td>
<td>Construct validity</td>
</tr>
<tr>
<td>CSI</td>
<td>Hwang et al</td>
<td>OBS</td>
<td>24 items</td>
<td>NR</td>
<td>$\alpha=0.84$</td>
<td>NR</td>
</tr>
<tr>
<td></td>
<td></td>
<td>OBS</td>
<td>24 items</td>
<td>NR</td>
<td>$\alpha=0.68-0.85$</td>
<td>NR</td>
</tr>
<tr>
<td></td>
<td>Utne et al</td>
<td>OBS</td>
<td>24 items</td>
<td>NR</td>
<td>$\alpha=0.88$</td>
<td>NR</td>
</tr>
<tr>
<td></td>
<td>Yoon et al</td>
<td>OBS</td>
<td>24 items</td>
<td>NR</td>
<td>$\alpha=0.63-0.85$</td>
<td>NR</td>
</tr>
<tr>
<td>CSI</td>
<td>Hwang et al</td>
<td>OBS</td>
<td>24 items</td>
<td>NR</td>
<td>$\alpha=0.84$</td>
<td>NR</td>
</tr>
<tr>
<td></td>
<td>Meyers and Gray</td>
<td>OBS</td>
<td>13 items</td>
<td>NR</td>
<td>$\alpha=0.86$</td>
<td>NR</td>
</tr>
<tr>
<td>CSI</td>
<td>Chan and Suen</td>
<td>MES</td>
<td>13 items</td>
<td>NR</td>
<td>$\alpha=0.91$</td>
<td>NR</td>
</tr>
<tr>
<td>FACQ-PC</td>
<td>Cooper et al</td>
<td>MES</td>
<td>26 items</td>
<td>Experts</td>
<td>$\alpha=0.73-0.86$</td>
<td>FA=Presence of 6 initial factors with eigenvalues exceeding 1, explaining 25, 14, 8, 7, 5, and 4% of the variance, respectively. Caregiver strain subscale the strongest correlation with subjective burden.</td>
</tr>
</tbody>
</table>

(Continued)
<table>
<thead>
<tr>
<th>Measure</th>
<th>Study</th>
<th>Type of study</th>
<th>No. of items</th>
<th>Psychometric information</th>
<th>Content Validity</th>
<th>Internal consistency</th>
<th>Construct validity</th>
<th>Reliability</th>
<th>Responsiveness</th>
<th>Acceptability and feasibility</th>
<th>Measure references cited by study</th>
<th>Original validation studies of measure (study population)</th>
</tr>
</thead>
<tbody>
<tr>
<td>FACT</td>
<td>Northouse et al&lt;sup&gt;14&lt;/sup&gt;</td>
<td>RCT</td>
<td>NR 17 items</td>
<td>Correlations with subscales MQOL ( r&lt;0.39 ), though psychological symptoms and support are significantly correlated to existential wellbeing ( p&lt;0.0015 ), ( p&lt;0.005 ).</td>
<td>NR</td>
<td>( \alpha=0.76–0.73 )</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>Northouse 2002&lt;sup&gt;27&lt;/sup&gt;</td>
<td>Cella 1993&lt;sup&gt;178&lt;/sup&gt; (cancer patients)</td>
</tr>
<tr>
<td>MQOL</td>
<td>Sherman et al&lt;sup&gt;10&lt;/sup&gt;</td>
<td>OBS 17 items</td>
<td>NR</td>
<td>Correlation with distress ( r=-0.245 ), ( p=0.04 )</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>Cooper 2006&lt;sup&gt;15&lt;/sup&gt;</td>
<td>Northouse 2002&lt;sup&gt;27&lt;/sup&gt; (cancer patients)</td>
</tr>
<tr>
<td>MBCBS</td>
<td>O’Hara et al&lt;sup&gt;13&lt;/sup&gt;</td>
<td>RCT 14 items</td>
<td>NR 14 items</td>
<td>Correlation HCQ-c -WHOQOL-Bref = Overall quality of life ( r=0.688 ), ( p&lt;0.01 ), Physical domain ( r=0.415 ), ( p&lt;0.01 ), psychological domain ( r=0.570 ), ( p&lt;0.01 ), social domain ( r=0.561 ), ( p&lt;0.01 ), environmental domain ( r=0.619 ), ( p&lt;0.01 ), and global spirituality ( r=0.639 ), ( p&lt;0.01 ).</td>
<td>NR</td>
<td>( \alpha=0.76–0.90 )</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>Cohen 1996&lt;sup&gt;179&lt;/sup&gt;, Cohen Mount 1996&lt;sup&gt;142&lt;/sup&gt;</td>
<td>Montgomery 1989&lt;sup&gt;161&lt;/sup&gt; (Caregivers of elderly persons)</td>
</tr>
<tr>
<td>WHO-QOL</td>
<td>Paiva et al&lt;sup&gt;12&lt;/sup&gt;</td>
<td>MES 26 items</td>
<td>NR</td>
<td>Correlation ( r=0.36 ). The 16 items predicted 55% of variance in global QOL and 53% in 7 domain scores. QOLLTI-F Total score predicted 43%.</td>
<td>NR</td>
<td>( \alpha=0.80 )</td>
<td>NR</td>
<td>All scores significantly different between days, with exception of financial. 2 sub measures limited by ceiling effect.</td>
<td>NR</td>
<td>NR</td>
<td>Berlim 2005&lt;sup&gt;182&lt;/sup&gt; (Brazilian version), Chachamovich 2007&lt;sup&gt;183&lt;/sup&gt; (WHOQOL-BREF)</td>
<td>WHOQOL Group 1999&lt;sup&gt;184&lt;/sup&gt; (general population)</td>
</tr>
<tr>
<td>QOLLTI-F</td>
<td>Cohen et al&lt;sup&gt;14&lt;/sup&gt;</td>
<td>MES 16 items</td>
<td>Expert panel</td>
<td>Correlation QOLLTI-F-HIS ( r=0.40 ) (p=0.000), explained variance 16.2%.</td>
<td>Testeret ( r=0.77–0.8 ).</td>
<td>( \alpha=0.86 )</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>Schur et al&lt;sup&gt;19&lt;/sup&gt;</td>
<td>Cohen 2006&lt;sup&gt;14&lt;/sup&gt; (Caregivers of palliative cancer patients)</td>
</tr>
<tr>
<td></td>
<td>Schur et al&lt;sup&gt;19&lt;/sup&gt;</td>
<td>MES (German version) 16 items</td>
<td>Group of various stakeholders (relatives, clinicians)</td>
<td>Correlation QOLLTI-F-HIS (integrative hope scale) ( r=0.40 ) (p=0.000), explained variance 16.2%.</td>
<td>Testeret ( r=0.92 ).</td>
<td>( \alpha=0.85 )</td>
<td>NR</td>
<td>All items showed a rate of missing responses below 5%</td>
<td>NR</td>
<td>NR</td>
<td>Schur et al&lt;sup&gt;19&lt;/sup&gt;</td>
<td>Cohen 2006&lt;sup&gt;14&lt;/sup&gt; (Caregivers of palliative cancer patients)</td>
</tr>
<tr>
<td>Measure</td>
<td>Study</td>
<td>Type of study</td>
<td>No. of items</td>
<td>Psychometric information</td>
<td>Measure references cited by study</td>
<td>Original validation studies of measure (study population)</td>
<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>
<td>---------------------------------------------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Content Validity</td>
<td>Internal consistency</td>
<td>Construct validity</td>
<td>Reliability</td>
<td>Responsiveness</td>
<td>Acceptability and feasibility</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>QOLLTI-F –HADS</td>
<td>Stress r=−0.41 (p=0.000), explained variance 16.26%. QOLLTI-F –HADS Anxiety r=−0.51 (p=0.000), explained variances 26%. QOLLTI-F –HADS Depression r=−0.52 (p=0.000), explained variances 26%. QOLLTI-F – Subjective burden r=−0.55 (p=0.000), explained variances 29.1%. FA= problematic cross-loading above 0.32 for items 1, 2, 3, 8, 11, 14 and 16. Scale showed a stable 4-factor structure, instead of initial 7.</td>
<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>QOLS</td>
<td>Grov et al</td>
<td>OBS</td>
<td>16 items</td>
<td>NR</td>
<td>α=0.88</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>'Burckhardt 2003185, Wahl 1998186, Flanagan 1978144</td>
<td></td>
<td></td>
</tr>
<tr>
<td>QOLI</td>
<td>Scott</td>
<td>OBS</td>
<td>36 items</td>
<td>NR</td>
<td>α=0.70</td>
<td>NR</td>
<td>Total reliability r=0.92–0.96</td>
<td>NR</td>
<td>NR</td>
<td>Ferrans 1992187, McMillan 1994184</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SF-36</td>
<td>Ringdal et al</td>
<td>OBS</td>
<td>36 items</td>
<td>NR</td>
<td>α=0.70–0.94</td>
<td>Correlations from subscales ranges r=0.50–0.89</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>Ware 1992185, Loce 1998188, Jenkinson 1997189, McHorney 1993190, Jenkinson 1993191, Ware 1992185, McHorney 1993190, McHorney 1994192</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>MES</td>
<td>35 items</td>
<td>NR</td>
<td>r= 0.52 to 0.78 for subscales with other QoL measures. Test–Retest r=0.68–0.93</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>Ware 1998193</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SF-12</td>
<td>Duggleby et al</td>
<td>OBS</td>
<td>12 items</td>
<td>NR</td>
<td>√ SF-36 r=0.95 and r=0.96</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>Ware 1998193</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SWED-QOL</td>
<td>Persson et al</td>
<td>OBS</td>
<td>64 items</td>
<td>NR</td>
<td>α=0.68–0.93</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>&quot;Brown, 2005194, Eriksson 2005195, Lindquist 2000196, Jonsson 1999197, Bedard 2001197</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ZBI</td>
<td>Bentley et al</td>
<td>OBS</td>
<td>NR</td>
<td>NR</td>
<td>α=0.87</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>Zarit 1990155 (Caregivers of dementia patients)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(Continued)
<table>
<thead>
<tr>
<th>Measure Study</th>
<th>Type of study</th>
<th>No. of items</th>
<th>Psychometric information</th>
<th>Measure references cited by study</th>
<th>Original validation studies of measure (study population)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brink et al53</td>
<td>OBS 12 items</td>
<td>CFA examined factor structure of 12 abridged items. Correlation 12 items–22 items r=0.92–0.97. Mean scores were high 19.89, compared to Bedard 2001197 =11.20, O’Rourke 2003198=8.29, Higginson 2010199=2.0</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>Higginson et al81</td>
<td>OBS 22 items</td>
<td>ZBI-12 p=0.95–0.97, ZBI-8 p=0.86–0.93, ZBI-7 p=0.90–0.95, ZBI-6 p=0.89–0.95, ZBI-4 p=0.88–0.92, ZBI-I p=0.63–0.78</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>Prigerson et al62</td>
<td>OBS 9 items</td>
<td>Mean caregiver burden score was 8.17 for with MDD versus 5.71 for without MDD (p&lt;0.002).</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
</tr>
</tbody>
</table>

*Measure not used in study, ¹Conference article, ²Grey literature, ³Non English, ⁴Review, NR=Not Reported, MES=Methodological study, OBS=Observational study, RCT=Randomized Controlled Trial.

BCOS=Bakas Caregiving Outcomes Scale, BDHI=Beck Depression Inventory, BI C=The Burden Index of Caregivers, BSFC=Burden Scale for Family Caregivers, CB=Caregiver Burden Inventory, CBS=Caregiver Burden Scale, CBS–EOLC=Caregiver’s Burden Scale in End-of-Life Care, CES-D=Center for Epidemiologic Studies Depression Scale, CFA=Confirmatory factor analysis, CG=Clinical Global Impression, CIS=Caregiver Impact Scale, CQOLI=Caregiver Quality of Life Index, CQOLI-C=Caregiver Quality of Life Index-Cancer, CQOLI-R=Caregiver Quality of Life Index-Revised, CRA=Caregiver Reaction Assessment, CSI=Caregiver Strain Index, ECOG=Eastern Cooperative Oncology Group, EFA=Exploratory factor analysis, ESRD=End stage renal disease, FA=Factor Analysis, FACCQF=Family Appraisal of Caring Questionnaire for Palliative Care, FACT=Functional Assessment of Cancer Therapy, HADS=Hospital Anxiety and Depression Scale, IHS=Integrative Hope Scale, HCQ-C=Holistic Comfort Questionnaire – Caregiver, ICC=Intra-class Correlation, ITCC=Corrected Item-Total Correlation, LASA=Linear Analogue Scale Assessment, MBCBS=Montgomery Borgatta Caregiver Burden Scale, MCSDS=Maislowe-Crowne Social Desirability Scale, MDD=Majoor depressive disorder, MOS-SS=Mos Medical Outcomes Study-Social Support, MQOL=McGill Quality of Life Questionnaire, MSPSS=Multidimensional Scale of Perceived Social Support, PSR=Performance Status Rating, PSS=Perceived Stress Scale, QOL=Quality of life, QOLI=Quality of Life Index, QOLITI-F=Quality of Life in Life Threatening Illness-Family Carer Version, QOLS=Quality of Life Scale, SF=Short Form, STA=State-Trait Anxiety Inventory, SWED-QOL=Swedish Health-Related QOL Survey, WHO-QOL=World Health Organization Quality of Life, ZBI=Zarit Burden Inventory.
these have been widely validated in a large number of different populations, it seems reasonable to assume that they are applicable for carers in a palliative context as well.

It is interesting that limited psychometric information was reported for the most widely used carer-specific measure, the Caregiver Reaction Assessment.\(^{165}\) This suggests that psychometric properties of the measures may not be the key factor in researchers’ choice of outcome measures. It would be worthwhile exploring in further studies what considerations researchers take into account when selecting their measures and why some carer-specific measures are used more frequently than others, particularly those developed specifically for carers in a palliative context.

Choosing the right measure for a particular study can be challenging because there may be a number of relevant measures from which to choose.\(^ {205}\) A systematic review would be appropriate valuable method to identify the most suitable measure, but it may not always be feasible to conduct a systematic review. Alternatively, as our systematic review highlights, no measure may seem entirely appropriate due to a lack of psychometric information. Additionally, measures may include items irrelevant to the study population, but developing new measures is costly and time consuming. Measure listings such as the Mapi research trust\(^ {206}\) and published systematic reviews can assist in selecting an appropriate measure.\(^ {205}\) Studies in this review did not always reference the measures used or when a reference was provided, it was frequently not the reference of the development of the measure. We encourage authors to reference the original development paper(s) of the measure(s) used and to justify their choice of instrument.

The findings of this systematic review are in line with previously published reviews. Hudson et al.\(^ {24}\) identified 62 tools covering a range of topics including satisfaction, experience, bereavement, needs, preparedness, family functioning and outcomes. Hudson et al.\(^ {24}\) identified a larger number of tools than we did as they included instruments, which we specifically excluded. The review concluded that appropriate tools were lacking but the authors only gave a broad critical appraisal across substantially different types of instruments. In 2009, Whalen and Buchholz\(^ {207}\) identified 74 caregiver burden screening tools for children or adults providing informal care, not specific to a palliative care context. Whalen and Buchholz\(^ {207}\) reported that burden measures might seem appropriate for informal carers but many are lacking psychometric information. Deeken et al.\(^ {208}\) searched MEDLINE and PubMed from 1966 to 2002 and identified 28 tools on burden (\(n = 17\)), needs (\(n = 8\)) and quality of life (\(n = 3\)). Neither Whalen and Buchholz nor the Deeken et al. reviews focused on palliative care. In contrast, our systematic review was conducted in a broader range of databases, specifically focused on self-reported multidimensional carer outcome measures in a palliative care context.

A strength of this systematic review is the comprehensive search of eight databases using four main search terms and no date restrictions, which meant we could collate and examine the variety of outcome measures that have been used with informal carers in a palliative care context. This review shows that although there is an increasing number of studies of informal carers in palliative care, most of the outcome measures used have not been formally validated within this carer population.

Another strength of the review is the care that was taken with regard to the inclusion criterion of palliative care. Palliative care is a complex process and involves a broad spectrum of health care services and treatments. Not all palliative care studies are labelled as such but refer to ‘hospice care’ or ‘end-of-life care’. These search terms were included but provided some challenges. For example, end-stage renal failure is for some patients a chronic disease but when dialysis or treatment is no longer effective, patients need a palliative approach. Two palliative care experts (A.A. and B.W.) independently assessed each study where there was uncertainty to determine whether or not it was in a palliative care population.

A limitation of the review is the exclusion of the grey literature and literature in languages other than English. It is likely that this meant we missed measures published outside the standard academic field or validation studies of translated measures, which might have provided further psychometric information.

A second limitation is rooted in the limitations of literature itself. Limited psychometric information was available, as more than half of the studies (\(n = 60\)) did not report any psychometric data. We included all studies that used multidimensional outcome measures in informal carers in palliative care, rather than only development or validation studies, as this corresponded to our study aims. We did not intend to include only development or validation studies, but this may be more appropriate for assessing psychometrics. However, if our inclusion criteria had been limited to development or validation studies alone, only four studies\(^ {34–36,94}\) would have been identified. Trends regarding the increasing number of publications on carer outcomes in palliative care would have been missed. As most of the studies did not include psychometric information, we could not critically assess the quality of most of the measures.

**Conclusion**

Support for patients receiving care is likely to continue to be devolved to informal carers. The WHO has called for health care provision to be extended to families, ensuring their needs, coping and outcomes are addressed alongside those of patients receiving health care services at the end of life.\(^ {209}\) As more interventions are developed to support carers, carers’ outcomes will increasingly be assessed in
palliative care context. Although a wide range of measures have already been used in this context, very limited formal psychometric testing has been undertaken. The frequently used measures contain limited psychometric information, while the outcome measures developed or validated in this context are not frequently used in research. Hence, further development and refinement of measures for informal carers in palliative care is required in order to be able to sufficiently support informal carers.

Acknowledgements
The authors would like to thank librarian Nia Roberts (Bodleian Health Care Libraries, University of Oxford, Oxford, UK) for helping us create the syntaxes for the electronic databases.

Declaration of conflicting interests
The author(s) declared no potential conflicts of interest with respect to the research, authorship and/or publication of this article.

Funding
The author(s) received no financial support for the research, authorship and/or publication of this article. Michele Peters is a senior researcher of the Department of Health funded Policy Research Unit on Quality and Outcomes of Person-centred Care, a collaboration between the London School of Economics and Political Science (LSE) and the Universities of Kent and Oxford.

References
5. Claxton-Olfield S, MacDonald J and Claxton-Olfield J. What palliative care volunteers would like to know about the patients they are being asked to support. Am J Hosp Palliat Care 2006; 23: 192–196.


96. Meyers JL and Gray LN. The relationships between family primary caregiver characteristics and satisfaction with


190. McHorney CA, Ware JE, Jr and Raczek AE. The MOS 36-Item Short-Form Health Survey (SF-36): II. Psychometric and clinical tests of validity in measuring physical and mental health constructs. Med Care 1993; 31: 247–263.


