

*Special Article*

# The Selection and Use of Outcome Measures in Palliative and End-of-Life Care Research: The MORECare International Consensus Workshop

Catherine J. Evans, BSc (Hon), MSc, PhD, RN, Hamid Benalia, BSc, MA, Nancy J. Preston, BSc, PhD, RN, Gunn Grande, BA, MSc, MPhil, PhD, Marjolein Gysels, BA, MA, PhD, Vicky Short, MPhil LLB (Hons), Barbara A. Daveson, BA, PhD, BMUS (MUSTHY), Claudia Bausewein, MSc, MD, PhD, Chris Todd, BA, MA, PhD, and Irene J. Higginson, BMBS, BMedSci, PhD, FFPHM, FRCP, on behalf of MORECare

*Department of Palliative Care, Policy and Rehabilitation (C.J.E., H.B., M.G., B.A.D., C.B., I.J.H.), Cicely Saunders Institute, King's College London, London; School of Nursing, Midwifery & Social Work (N.J.P., G.G., V.S., C.T.), University of Manchester, Manchester; International Observatory on End of Life Care and Faculty of Health & Medicine (N.J.P.), Lancaster University, Lancaster, United Kingdom; Centre for Social Science and Global Health (M.G.), University of Amsterdam, Amsterdam, The Netherlands; and Interdisciplinary Centre for Palliative Medicine (C.B.), Munich University Hospital, Munich, Germany*

---

## Abstract

**Context.** A major barrier to widening and sustaining palliative care service provision is the requirement for better selection and use of outcome measures. Service commissioning is increasingly based on patient, carer, and service outcomes as opposed to service activity.

**Objectives.** To generate recommendations and consensus for research in palliative and end-of-life care on the properties of the best outcome measures, enhancing the validity of proxy-reported data and optimal data collection time points.

**Methods.** An international expert “workshop” was convened and an online consensus survey was undertaken using the MORECare Transparent Expert Consultation to generate recommendations and level of agreement. We focused on three areas: 1) measurement properties, 2) use of proxies, and 3) measurement timing. Data analysis comprised descriptive analysis of aggregate scores and collation of narrative comments.

**Results.** There were 31 workshop attendees; 29 recommendations were included in the online survey, completed by 28 experts. The top three

---

*Address correspondence to:* Catherine J. Evans, PhD, MSc, BSc, RN, Department of Palliative Care, Policy and Rehabilitation, Cicely Saunders Institute, King's

---

College London, Bessemer Road, London SE5 9PJ, United Kingdom. E-mail: [catherine.evans@kcl.ac.uk](mailto:catherine.evans@kcl.ac.uk)

*Accepted for publication:* January 28, 2013.

recommendations by area were the following: 1) the properties of the best outcome measures are responsive to change over time and capture clinically important data, 2) to enhance the validity of proxy data requires clear and specific guidelines to aid lay individuals' and/or professionals' completion of proxy measures, and 3) data collection time points need clear identification to establish a baseline.

**Conclusion.** Outcome measurement in palliative and end-of-life care requires the use of psychometrically robust measures that are clinically responsive, with defined data collection time points to establish a baseline and clear administration guidelines to complete proxy measures. To further the field requires clinical imperatives to more closely inform recommendations on outcome measurement. *J Pain Symptom Manage* 2013;46:925–937. © 2013 U.S. Cancer Pain Relief Committee. Published by Elsevier Inc. All rights reserved.

### **Key Words**

*Outcome assessment, evaluation studies, research design, palliative care, consensus*

---

## **Introduction**

Widening access to palliative and end-of-life (EOL) care services is advocated with corroboration of patient<sup>1</sup> and carer benefits,<sup>2</sup> greater potential for health service cost savings,<sup>3,4</sup> and increasing demand with an aging population.<sup>5</sup> A major barrier is the requirement for better selection and use of outcome measures to demonstrate the effectiveness of services. Commissioners of health and social care services increasingly require service providers to use patient, carer, and service outcomes to demonstrate a service's safety, effectiveness, and quality as opposed to detailing service activity.<sup>6,7</sup> In research and clinical practice, a multitude of measures are used, and frequently these measures are not validated with palliative care populations.<sup>8</sup> This hampers meta-analyses, limits responsiveness to change in outcome, evaluating service effectiveness, and developing the evidence base to inform best practice.<sup>9</sup>

Trials and nonrandomized designs in palliative and EOL care are often compromised by the use of untested outcome measures,<sup>10,11</sup> measures not developed for palliative care populations,<sup>12</sup> uncertainty as to the best measurement time points,<sup>13</sup> and the use of measures for symptom change with less use of measures encompassing the multiple domains of palliative care (e.g., Palliative care Outcome Scale [POS], Edmonton Symptom Assessment System).<sup>12</sup> These limitations are not unique to palliative and EOL care. The COnsensus-based Standards for the selection of health

Measurement INstruments is a comprehensive checklist for assessing the quality of the measurement properties of health status questionnaires. The checklist was developed in response to the rapid increase in health status questionnaires and the need for quality criteria to compare measures in systematic reviews, identify shortcomings, and design studies validating measures.<sup>14,15</sup>

Palliative and EOL care involves people with increasing debility associated with advancing disease and has a broad mandate of care provision, with intended outcomes of improving quality of life for patients and their caregivers.<sup>16</sup> The nature of palliative and EOL care requires measurement properties that accommodate the following: the multiple domains of palliative care; to change over time and increasing levels of fatigue; the use of proxies, particularly when individuals are near to death; and timing to detect change and monitor sustainment of change. These challenges are well reported, and international advancements to address them are evident: projects such as the Palliative Care Outcomes Collaboration in Australia that uses national standardized assessments of palliative care outcomes<sup>17,18</sup> and a European collaboration, entitled PRISMA, focused on promoting best practice in the measurement of EOL care (Preferred Reporting Items for Systematic Reviews and Meta-Analyses)<sup>11</sup> working with a European Association of Palliative Care (EAPC) Taskforce on patient-reported outcome measures (PROMs) in palliative care to

further the work<sup>19,20</sup> and provide guidance on selecting and using measures.<sup>21</sup> Internationally, there is increasing use of multi-domain palliative care measurement tools in research and clinical practice (e.g., POS and Support Team Assessment Schedule [STAS]).<sup>22</sup>

No work, however, has considered specific recommendations for the use and selection of outcome measures in trials and nonrandomized studies of palliative and EOL care interventions. This article reports on an international outcome measurement workshop and online consensus survey that aimed to identify agreed best practice on selecting and using outcome measures. The work is part of the Methods Of Researching End of life Care (MORECare) project, developing evidence-based guidance on the design and conduct of evaluative research on palliative and EOL care. MORECare integrated data from three systematic reviews<sup>23–25</sup> and five workshops<sup>26,27</sup> to generate a statement (the MORECare statement) that sets clear standards on good research practice in evaluating services and treatments in palliative and EOL care.<sup>28</sup> This paper reports on the outcome measurement workshop.

## Methods

### Setting

An invitation-only workshop was held as a joint initiative between the MORECare and

PRISMA projects at the 2011 EAPC conference to further the European work undertaken by PRISMA<sup>11</sup> and the EAPC Taskforce on PROMs in palliative care.<sup>19</sup>

### Participants

Holding the workshop at an international conference intended to enable participation by international experts in palliative care research, service provision, and policy. Participants were identified from the EAPC online conference program and considered experts from their publications. Fifty-nine individuals were identified and received a personal e-mail invitation to attend the workshop and a single reminder e-mail.

### Design

We used the MORECare Transparent Expert Consultation method, which incorporates consensus methods of nominal group and consensus survey.<sup>29</sup> The Transparent Expert Consultation comprised a staged process: Stage I—identifying critical issues from the literature; Stage II—a workshop with experts on the issues using nominal group technique to generate recommendations; and Stage III—an online consensus survey to prioritize and identify areas of contention/uncertainty (Fig. 1).

*Stage I: Identifying Critical Issues.* Critical issues on selecting and using outcome measures in

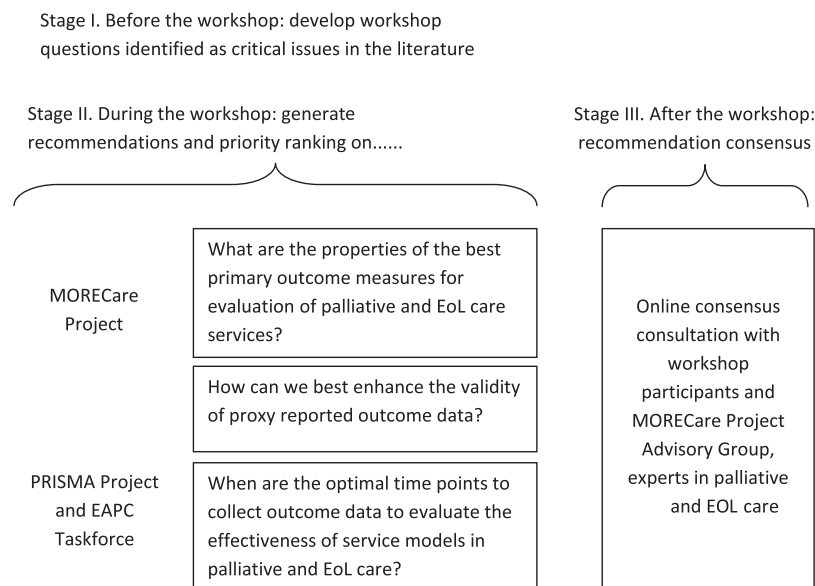


Fig. 1. Flow diagram of the study design.

palliative and EOL care were identified from systematic reviews of intervention studies<sup>12</sup> and outcome measurement.<sup>30</sup> The workshop and consensus survey focused on three critical questions:

1. What are the properties of the best primary outcome measures for evaluation of palliative and EOL care services?
2. How can we best enhance the validity of proxy-reported outcome data?
3. When are the optimal time points to collect outcome data to evaluate the effectiveness of service models in palliative and EOL care?

*Stage II: Workshop.* Participants received a pre-workshop briefing pack detailing the aim, critical questions, and workshop format. The workshop format comprised presentations on the MORECare project and the project's interface with the PRISMA/EAPC Taskforce, followed by three structured groups of seven to 13 participants each focused on one of the critical questions. The group structure used a modified nominal group facilitated and scribed by members of the MORECare and PRISMA projects. The facilitators guided participants through a structured process of: 1) brief discussion, 2) individual writing of recommendations and ranking, and 3) individuals in turn reading out their highest ranked recommendations until individual lists were exhausted (or time exceeded).<sup>31</sup> The scribes wrote the recommendations and ranking on flip-charts, and each small group discussed and agreed on the final priority order, which they presented and discussed with the whole group. Participants listed and ranked recommendations from one to five (highest to lowest) on structured A4 sheets detailing the respective group question, the ranking scale

and boxes to list recommendations, rank, and detail rationale.

*Stage III: Consensus Survey.* Recommendations were posted online to the workshop participants (excluding the facilitators and scribes) and members of the MORECare Project Advisory Group. Participants received a personalized e-mail invitation and reminder after two weeks. The online participants ranked, from one to nine (strongly disagree to strongly agree), the extent they agreed with a recommendation and used free-text spaces to comment on each recommendation.

#### Data Analysis

Individual recommendations and ranking were entered into Excel spreadsheets with assigned participant identification numbers. Flip-chart records and the scribes' notes were typed. Three researchers (C.J.E., H.B., M.G.) drew on qualitative analysis to identify themes in the recommendations and assign codes<sup>32</sup> and collate typed data. Coded recommendations were rearranged by theme in Excel spreadsheets. Duplicates were combined and recommendations arranged by priority ranking. The recommendations retained participants' original language, with amendments to enhance clarity. A third researcher (I.J.H.) reviewed the analysis and proposed recommendations. After combining for duplicates, we included all recommendations in the online survey that participants had ranked three and above and/or participants' listed recommendations formed a prominent theme. The MORECare research team reviewed and agreed on the recommendations for the online survey.

Analysis of scaled data followed conventional rules of descriptive statistics (frequencies and medians) and plots (box and whisker plots of interquartile ranges to

Table 1  
Recommendation Agreement: Interpretation of Median Regions and IQR<sup>34</sup>

Median Regions and IQR	Interpretation
7–9	Recommendations are indicated
4–6	Recommendations are equivocal
1–3	Recommendations not indicated
IQR in <i>one</i> region	Strict agreement for recommendation
IQR in <i>any</i> three-point region	Broad agreement for recommendation

IQR = interquartile range.

analyze<sup>33</sup> and interpret levels of agreement [Table 1].<sup>34</sup> Narrative comments were collated by recommendation, themes identified, and rearranged in Excel sheets to enable comparison.<sup>35</sup> The analysis of comments intended to aid understanding and provide illustrative examples of the issues raised by the proposed recommendations.<sup>36</sup>

### *Ethics*

The study was approved by the University of Manchester Research Ethics Committee (reference number: 10328).

## **Results**

### *Participants*

Thirty-one international and national experts attended the workshop comprising senior academics, policy makers, and clinicians, representing countries from Europe (e.g., Norway, Belgium, U.K., Italy, Germany), North America, and Australia. All held a title of professor or doctor. Twenty-eight experts completed the online survey.

### *Recommendations and Levels of Agreement*

The workshop generated 155 individual recommendations; 29 recommendations were included in the online consensus survey after combining duplicates and completing analysis. The recommendations included in the online survey were either generic describing good practice in choosing and administering outcome measures in effectiveness research or focused on specific challenges in research on palliative and EOL care (Table 2). Overall, the median levels of agreement for the 29 recommendations showed that most recommendations ( $n = 26$ ) were indicated with medians  $\geq 7$ ; none were not indicated (medians  $\leq 3$ ) (Table 2). However, the levels of agreement across the three areas varied. The recommendations on: 1) the properties of outcome measures showed mainly strict agreement that the recommendations were indicated ( $n = 6/11$  [Fig. 2]), 2) enhancing proxy data showed mainly broad agreement ( $n = 6/9$  [Fig. 3]), and 3) data collection time points all showed broad agreement ( $n = 9$  [Fig. 4]). Recommendations in which the level of agreement was equivocal or broad

were considered as uncertain or contentious. Qualitative analysis of the workshop discussions and the online survey free-text comments informed understanding on these areas. We present quotes to illustrate these debates.

### *Properties of the Best Outcome Measurements*

Properties of the best primary outcome measures formed 11 recommendations (Fig. 2; Table 2). Overall, there was strict agreement that the properties of the best outcome measures required valid measures that were responsive to change over time, captured clinically relevant data, and were easy to administer across care settings (Recommendations 1, 2, 5, 6, 9, 10). The text data indicated that measures required flexibility in the administration of the measure (e.g., PROMs, observation) to ensure responsiveness to patients' fluctuating capacity and increasing fatigue, notably in the last days and hours of life. An area of less certainty was the recommendation that a measure was appropriate for clinical practice, research, and audit, with broad agreement indicated but with the widest spread (Recommendation 11) (Fig. 2).

Uncertainty centered on accommodating competing priorities for practice, research, and audit, summarized by a survey respondent as:

It is impossible to serve all 3 masters optimally. They need to be optimized for one or perhaps two goals. (ID OMs 6)

Measures in research required greater detail than the comparatively brief measures required in clinical practice. However, there was agreement on the importance of research capturing clinically relevant data (Recommendation 5) that could be integrated into clinical care (Recommendation 10). A way forward discussed in the workshop and in the online survey comments was the use of "core outcomes" with additions depending on the context, as a survey respondent stated:

... core outcomes (for all settings) and others setting-specific, this will give a more accurate approach, for the process and for specific settings. (ID OMs 12)

### *Enhancing the Validity of Proxy-Reported Outcome Data*

Greater uncertainty surrounded the seven draft recommendations on enhancing the

Table 2  
**Recommendations on Selecting and Using Outcome Measures by Area and Level of Agreement**

Area 1: Properties of the Best Primary Outcome Measures in Evaluations of Palliative and EOL Care		
No.	Should ...	Median (IQR)
<b>Strict agreement that recommendation is indicated</b>		
1	Be easy to administer and interpret (e.g., short and low level of complexity).	8 (7–9)
2	Be applicable across care settings to capture change in outcomes by location (e.g., patient's home, hospital, hospice).	8 (7–9)
5	Be responsive to change over time and capture clinically important data.	8 (7.5–8)
6	Have demonstrated content validity.	8 (7–9)
9	Have demonstrated reliability.	8 (7–9)
10	Be integrated into clinical care.	8 (7–9)
<b>Broad agreement that recommendation is indicated</b>		
3	Work across a disease trajectory from diagnosis to death.	7 (6–8)
4	Be culturally sensitive to a respective population group (e.g., linguistically and culturally).	7.5 (6–8)
7	Have demonstrated construct validity.	8 (6–9)
8	Have demonstrated face validity.	8 (6–9)
11	Be appropriate for clinical practice, research, and audit uses.	7 (6–9)
Area 2: Enhancing the Validity of Proxy Data in Evaluations of Palliative and EOL Care Services		
No.	Requires ...	Median (IQR)
<b>Strict agreement that recommendation is indicated</b>		
14	Clear and specific guidelines to aid lay individuals' and/or professionals' completion of proxy measures.	8.5 (7–9)
20	Capture of patients' experiences of care (e.g., emotional effect, physical effect).	8 (7–9)
17	That data from patients and proxy measures should be differentiated in the dataset.	8 (7–9)
<b>Broad agreement that recommendation is indicated</b>		
12	PROMs as the gold standard in evaluations of palliative and EOL care services; proxy measures should be used only when the patient concerned is unable to provide the data.	7 (5–9)
15	Proxy measures to have demonstrated criterion validity.	7 (6–9)
16	Proxy measures to have demonstrated interrater validity.	7 (6–9)
19	Capture of the core intervention.	7 (6–9)
<b>Broad agreement that recommendation is equivocal</b>		
13	That to increase reliability in evaluations of palliative and EOL care services, proxy measures should only be used in conjunction with observer measures.	6 (5–7)
18	Factors that may influence proxy raters to be investigated and reported (e.g., degree of emotional involvement or neutrality).	6.5 (4–7.3)
Area 3: Data Collection Time Points in Evaluations of Palliative and EOL Care Require ...		
No.	Require ...	Median (IQR)
<b>Broad agreement that recommendation is indicated</b>		
21	Clear identification of time points to establish a baseline.	7 (6–9)
22	Time points in evaluations on palliative and EOL care services need to be established before conducting the evaluation.	7 (5–8)
23	That when prospective measurement is used, endpoints should correspond to when the effect of the intervention is expected to take place.	7 (6–8)
24	That when prospective measurement is used, time points should correspond to a combination of fixed and flexible time points (e.g., a flexible time point may be triggered when a threshold is reached).	7 (6–8)
26	Process measures to understand the implementation of the intervention.	8 (6–9)
27	A measurement window for time points is defined to capture baseline data and the impact of the intervention.	7 (5–7.3)
28	The choice of time points is influenced by the expected effect of the intervention, the study setting(s) (e.g., patient's home, hospital) and the disease trajectory and stage.	7 (5.8–8.3)
29	That the frequency of measurement is determined by the number, length, and degree of burden of the data collection tools.	7 (6–8)
<b>Broad agreement that recommendation is equivocal</b>		
25	In retrospective evaluations that death should form the endpoint.	6 (4–7)

EOL = end of life; IQR = interquartile range; PROMs = patient-reported outcome measures.

validity of proxy-reported data, with wide variance in degrees of agreement (Fig. 3). There was strict agreement that to enhance proxy measures required clear specific guidelines to aid completion, the differentiation of patient and proxy data in datasets and that proxy

measures captured patients' experiences of care (Recommendations 14, 17, 20) (Table 2).

Two recommendations were equivocal and concerned increasing the reliability of proxy measures and understanding bias (Recommendations 13 and 18). Recommendation 13 that

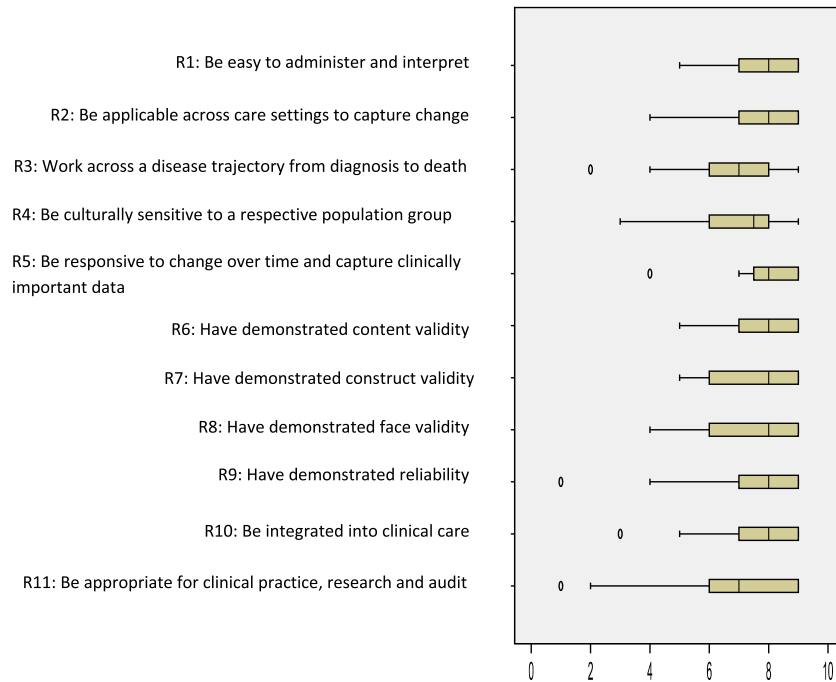


Fig. 2. Properties of the best primary outcome measures in evaluations of palliative and EOL care. Box and whisker plot of the interquartile ranges and medians of level of agreement for the 11 recommendations (box: 25th and 75th percentiles; whiskers: minimum and maximum).

proxy measures should only be used in conjunction with observer measures was viewed as useful in the absence of PROMs data. However,

uncertainty surrounded the validity of using such measures in conjunction and the best method of statistical analysis. Best research

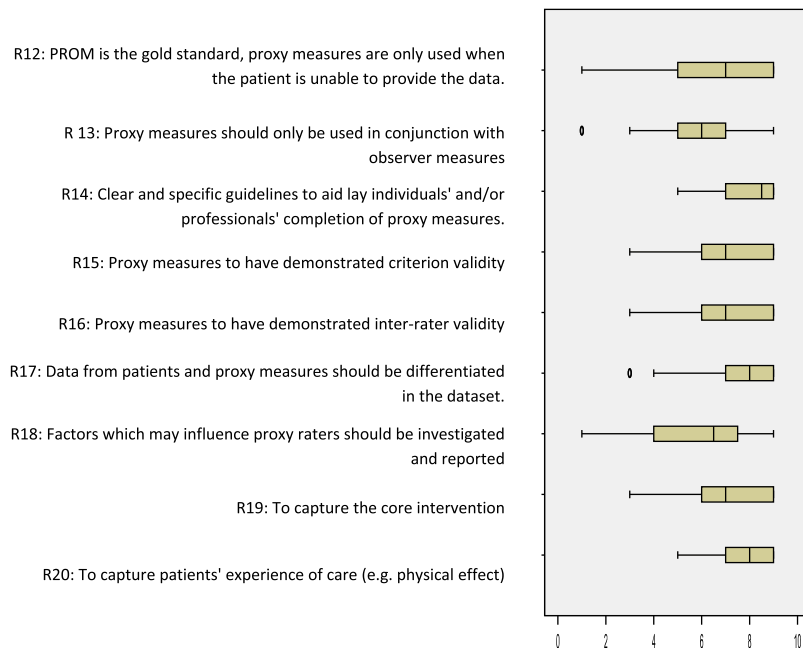


Fig. 3. Enhancing the validity of proxy data in evaluations of palliative and EOL care. Box and whisker plot of the interquartile ranges and medians of level of agreement for the nine measurement recommendations (box: 25th and 75th percentiles; whiskers: minimum and maximum).

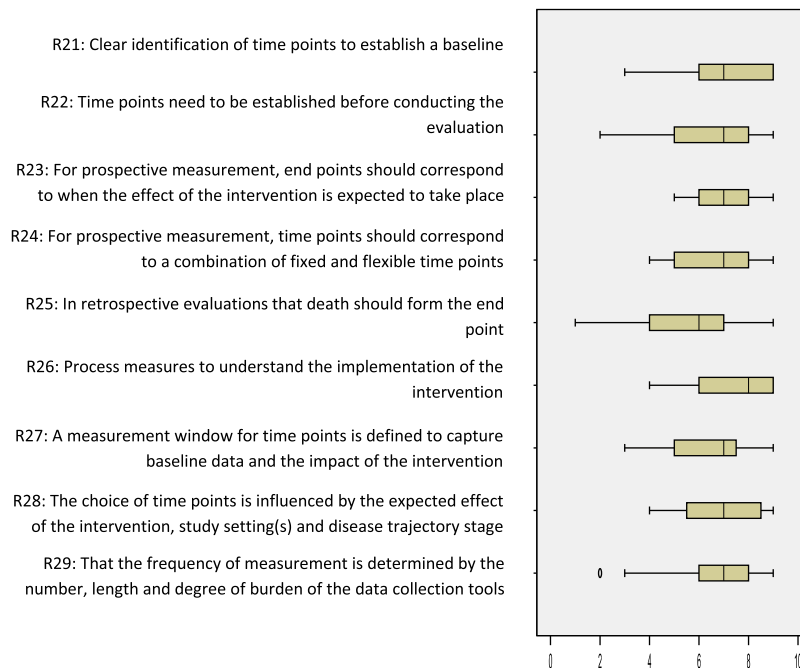


Fig. 4. Data collection time points in evaluations of palliative and EOL care. Box and whisker plot of the inter-quartile ranges and medians of level of agreement for the nine recommendations (box: 25th and 75th percentiles; whiskers: minimum and maximum).

practice meant using valid measures, but further research was required on the contribution of using measures in conjunction to enhance reliability. There was agreement that Recommendation 18 was indicated on the importance of investigating factors that may influence proxy raters to reduce bias. However, participants' comments questioned the feasibility of this in research practice because of the complexity in collecting and interpreting such data.

The main area of contention was the recommendation on the precedence of PROMs over proxy measures, indicated by the wide spread (Recommendation 12) (Fig. 3). The survey participants' comments indicated that using PROMs data alone negated that palliative care encompassed both patients and their families, would fail to capture data in the last days of life from, for example, observation, and the relevance of clinical data, for example, medication use. Survey participants advocated the requirement for multiple measures in research on palliative care. The workshop delegates advocated that proxy measures were "second choice" over PROMs, but proxy data were "indispensable" in palliative care because of patients' deteriorating capacity and increasing fatigue.

#### *Optimal Data Collection Time Points in Clinical Trials and Evaluations of Interventions*

There was broad agreement that the eight recommendations on time points were indicated (Fig. 4). The wide distribution and no strict agreement on the recommendations suggested uncertainty/contention. The recommendation of death as the endpoint in retrospective data analysis was equivocal (Recommendation 25). The online survey comments revealed that death as an endpoint was considered too narrow as it failed to capture the holistic nature of palliative care, which encompasses carers and family members across the disease trajectory and into bereavement. However, death could form an important endpoint in retrospective data analysis depending on the objectives of the study, as one workshop participant stated:

Use death as [an] endpoint and count burden from this – as time or referral/first assessment may vary so greatly. (ID G2.015)

Although the workshop participants used the term endpoint, this caused confusion in the online survey, and alternative terms of anchor point or starting point were proposed.



The confusion likely contributed to wide range of scores and finding of an equivocal recommendation.

## Discussion

Research on palliative and EOL care is beset by well-reported challenges rooted in involving individuals with advanced illness in research and their families, the multiple domains of palliative care provided across health and social care, and the intention to improve quality of life in the face of deterioration.<sup>9,10,12,37</sup> These challenges have hampered the use and selection of outcome measures in trials and nonrandomized studies on palliative and EOL care services.<sup>38,39</sup> We critically discuss three main areas to further this field of research: priority measurement properties, incorporating proxy data, and identifying time points.

### Measurement Properties

The findings corroborate work undertaken by the PRISMA project that research on palliative care requires the use of short, brief outcome measures that accommodate patients' increasing levels of fatigue and minimize burden.<sup>8,20,40</sup> This does not equate to "easy" to measure aspects of care, but areas important to patients and families that encompass desired benefit from a service or intervention.<sup>41</sup> The challenge is to use measures that capture the multiple domains of palliative care, but in a format that is short and easy to interpret. Our findings support the requirement for measures used in palliative care to be applicable for both clinical practice and research.<sup>11,41</sup> Work in Australia has demonstrated the feasibility of collecting national palliative care outcome data across specialist services using defined measurement tools incorporated in routine practice and the potential for national benchmarking for care quality and outcomes.<sup>17</sup> However, our findings and the PRISMA work show that there are competing priorities between the requirements of a measure used in clinical practice and research.<sup>40</sup> Time pressure in clinical practice necessitates the use of short measures (e.g., six to 10 items)<sup>8</sup> that are easy to administer and interpret to inform clinical decisions. In research

evaluating the effectiveness of an intervention or service, the measure(s) used needs to demonstrate the degree of change over time, not simply the presence or absence of, for example, a symptom. Both clinical and research settings, however, require the use of measures with validated psychometric properties to ensure scientific rigor and enhance clinical decision making.<sup>40</sup> To accommodate these competing priorities (and similarities) of our findings, the PRISMA work suggests the use of core measures for clinical and research practice and further assessment of specific areas for research when indicated.<sup>8</sup> Mularski et al.<sup>42</sup> propose, from consensus work, that evaluation studies use an outcomes measurement strategy that encompasses three outcome categories: condition specific, patient outcomes, and family/caregiver outcomes. Measures such as POS (<http://pos-pal.org>),<sup>43</sup> the Edmonton Symptom Assessment System,<sup>44</sup> and Memorial Symptom Assessment Scale<sup>45</sup> are examples of measures commonly used in both palliative care clinical practice and research,<sup>41</sup> are validated with palliative care populations, and are short to administer and multidimensional.

Palliative care is provided internationally across health and social care settings and to patients with cancer and increasingly to those with nonmalignant illness. Our data indicate the requirement for measures to be culturally sensitive and be applicable across care settings to capture change in outcomes, particularly for patients who move between care settings, for example, at home, inpatient hospital, or care home. The development of existing measures is apparent, notably POS, with work to increase cultural sensitivity for countries both in Africa<sup>46</sup> and Europe,<sup>47</sup> and use with disease groups other than cancer.<sup>22</sup> Further validation work of palliative care measures like POS, however, is required for specific patient groups, notably patients with dementia<sup>48,49</sup> and for social care settings (e.g., care homes).<sup>50</sup>

### Incorporating Proxy Data

Our data showed the clinical complexity for palliative care in using PROMs as the "gold standard" in research. The complexity of palliative care service provision to meet diverse care needs of patients and families necessitates data collection from multiple sources. It is

an essential requirement in research on palliative and EOL care to expect and plan for missing data because of deterioration.<sup>27,42</sup> Proxy-reported outcome data are integral sources in palliative care, particularly at points of a patient's deterioration or unstable symptoms when physical and/or mental debility may prevent use of a PROM. At these points, patients likely experience increasing symptom distress, points at which it is imperative to understand the effectiveness of the care and treatment provided. PROMs, such as POS and Memorial Symptom Assessment Scale, are validated for use with a proxy.<sup>43,49,51</sup>

However, the extent a proxy report reflects a patient response depends on the degree of agreement and the direction of agreement (e.g., proxies report high ratings when patients report low) between the proxy and patient; the greater the level of agreement, the greater the confidence in including proxy reports.<sup>51</sup> Studies in palliative care have shown wide variance in agreement responses depending in particular on the variable measured and the patient-proxy dyad.<sup>51–53</sup> Studies, for example, report high agreement for observable patient variables (e.g., ability to transfer) and moderate to low for psychological domains and subjective patient experiences (e.g., quality of life).<sup>53,54</sup> In patient and caregiver dyads, higher agreement is reported than in patients and health care provider dyads,<sup>51</sup> but over time, similar reporting is seen as practitioners' knowledge of their patients increases.<sup>53</sup> Our findings correlate with studies on proxy and patient reporting that to further the inclusion of proxy reports requires greater understanding of the factors that influence reporting to improve reliability (e.g., degree of neutrality, knowledge of patient), further development of proxy measures to capture patients' subjective experiences (e.g., quality of life), and ways to enhance statistical analysis.<sup>51–54</sup>

### *Identifying Time Points*

Identifying optimal time points intends to maximize the number of patients available at consecutive points of measurement to demonstrate the effectiveness of the intervention.<sup>13</sup> Our data show that best practice to identify time points was the area of most uncertainty, with broad or equivocal agreement surrounding the recommendations posed. Identifying

time points in research on palliative and EOL care is a prominent methodological challenge rooted in minimizing patient burden, imprecision in identifying the state of "dying," and attrition.<sup>13,34</sup> In particular, attrition from increasing debility minimizes data collection for individuals with the greatest symptom burden and compromises a study's internal validity with overrepresentation of the "healthiest" patients.<sup>55</sup> Hence, careful planning of time points is crucial.

Our data show broad agreement for the recommendations on establishing time points before the evaluation, and the frequency of measurement is determined by the degree of burden of the data collection tools. These recommendations are likely too narrow, hence the wide spread of responses. The timing and frequency of measurements is embedded within a conceptual framework for the intervention. The conceptual framework is developed from the research question and detailed understanding of the intended intervention outcome/benefit, the expected response pattern (e.g., two weeks post-consultation with the palliative care clinician), and the phase of illness (e.g., stable disease, deteriorating, dying phase).<sup>18,56,57</sup>

The recommendation of death as an endpoint in retrospective evaluations was equivocal. This likely reflects misunderstanding of the term "endpoint." The use of anchor point from which data are included may have increased clarity. Retrospective evaluations in palliative care are important study designs. Prospective designs are hampered by imprecision in the identification of the state of dying, particularly in measures of quality of life, which often deteriorate in the last days of life.<sup>13</sup> A way to accommodate this is prospective observations and data analysis that uses both prospective analysis from the baseline and retrospective analysis from the last observation before death.<sup>58</sup> Death is a fixed anchor point for the whole sample and provides more reliable data on the clinical outcomes over time.<sup>59</sup>

### *Limitations*

We undertook the consensus workshop at an international conference in palliative care. This enabled participation by international experts but limited invitations to those attending

the conference and the time available, particularly for discussion. The use of clear questions and a structured process to generate the recommendations was intended to maximize the time. We limited the online consensus to a single round to enhance the response rate. This prevented wider consultation on the survey comments that detailed reasons for scoring as well as measures of central tendency and dispersion. There is no agreement on the best method for mathematical aggregation in consensus surveys.<sup>29</sup> We used an implicit approach of medians and measure of dispersion<sup>29</sup> and the Jones and Hunter established method for interpreting scaled data.<sup>30</sup>

### Conclusions

To enable the wide application and interpretation of outcome measures in palliative and EOL care, the measurement properties need to be validated for this population, capture multidimensional components, and be broadly applicable across health systems, populations, disease trajectories (including post-death for bereaved carers), languages, and cultures. The best measures have essential properties of simplicity of use and good interpretability, the ability to measure clinically important change, and reporting by a patient or proxy informant. To further the field requires clinical imperatives to more closely inform recommendations on outcome measurement.

### Disclosures and Acknowledgments

MORECare is funded by the National Institute for Health Research and managed by the Medical Research Council as part of the Methodology Research Programme (project number: G0802654/1). The authors declare no conflicts of interest.

MORECare aimed to identify, appraise, and synthesize “best practice” methods to develop and evaluate palliative and end-of-life care, particularly focusing on complex service delivery interventions and reconfigurations. Principal investigator: Irene J. Higginson; co-principal investigator: Chris Todd; co-investigators: Peter Fayers, Gunn Grande, Richard Harding, Matthew Hotopf, Penney Lewis, Paul McCrone, Scott Murray, and Myfanwy Morgan; project

advisory group: Massimo Costantini, Steve Dewar, John Ellershaw, Claire Henry, William Hollingworth, Philip Hurst, Tessa Inge, Jane Maher, Irene McGill, Elizabeth Murray, Ann Netten, Sheila Payne, Roland Petchey, Wendy Prentice, Deborah Tanner, and Celia A. Taylor; researchers: Hamid Benalia, Catherine J. Evans, Marjolein Gysels, Nancy J. Preston, and Vicky Short. We would like to thank the participants of the MORECare international outcome measurement workshop.

### References

1. Temel JS, Greer JA, Muzikansky A, et al. Early palliative care for patients with metastatic non-small cell lung cancer. *N Engl J Med* 2010;363:733–742.
2. Higginson IJ, Costantini M, Silber E, Burman R, Edmonds P. Evaluation of a new model of short-term palliative care for people severely affected with multiple sclerosis: a randomised fast-track trial to test timing of referral and how long the effect is maintained. *Postgrad Med J* 2011;87:769–775.
3. Morrison SR, Penrod J, Cassel B, et al. Cost savings associated with US hospital palliative care consultation programs. *Arch Intern Med* 2008;168:1783–1790.
4. Hatziaandreu E, Archontakis F, Daly A. The potential cost savings of greater use of home- and hospice-based end of life care in England. London: National Audit Office, 2008.
5. Gomes B, Higginson IJ. Where people die (1974-2030): past trends, future projections and implications for care. *Palliat Med* 2008;22:33–41.
6. Secretary of State for Health. Equity and excellence: Liberating the NHS. London: Department of Health, 2010.
7. Hughes-Hallet T, Craft A, Davies C, Mackay I, Nielsson T. Palliative care funding review: Funding the right care and support for everyone. London: Secretary of State for Health, Department of Health, 2011.
8. Harding R, Simon ST, Benalia H, et al. The PRISMA symposium 1: outcome tool use. Disharmony in European outcomes research for palliative and advanced disease care: too many tools in practice. *J Pain Symptom Manage* 2011;42:493–500.
9. Higginson IJ, Finlay IG, Goodwin DM, et al. Is there evidence that palliative care teams alter end-of-life experiences of patients and their caregivers? *J Pain Symptom Manage* 2003;25:150–168.
10. Hall S, Kolliakou A, Petkova H, Froggatt K, Higginson IJ. Interventions for improving palliative care for older people living in nursing care homes. *Cochrane Database Syst Rev* 2011;3:CD007132.

11. Harding R, Higginson IJ. PRISMA: a pan-European co-ordinating action to advance the science in end-of-life cancer care. *Eur J Cancer* 2010; 46:1493–1501.
12. Zimmermann C, Riechelmann R, Krzyzanowska M, Rodin G, Tannock I. Effectiveness of specialized palliative care: a systematic review. *JAMA* 2008;299:1698–1709.
13. Tang ST, McCorkle R. Appropriate time frames for data collection in quality of life research among cancer patients at the end of life. *Qual Life Res* 2002;11:145–155.
14. Terwee CB, Bot SDM, de Boer MR, et al. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol* 2007;60:34–42.
15. Terwee CB, Mokkink LB, Knol DL, et al. Rating the methodological quality in systematic reviews of studies on measurement properties: a scoring system for the COSMIN checklist. *Qual Life Res* 2012;21:651–657.
16. World Health Organization. National cancer control programmes: Policies and managerial guidelines. Geneva: World Health Organization, 2002.
17. Currow DC, Eagar K, Aoun S, et al. Is it feasible and desirable to collect voluntarily quality and outcome data nationally in palliative oncology care? *J Clin Oncol* 2008;26:3853–3859.
18. Eagar K, Watters P, Currow DC, Aoun SM, Yates P. The Australian Palliative Care Outcomes Collaboration (PCOC)—measuring the quality and outcomes of palliative care on a routine basis. *Aust Health Rev* 2010;34:186–192.
19. European Association for Palliative Care. EAPC taskforce on patient-reported outcome measurement in palliative care. Available from <http://www.eapcnet.eu/Themes/Clinicalcare/Outcomemeasurement.aspx>. Accessed July 9, 2012.
20. Bausewein C, Simon ST, Benalia H, et al. Implementing patient-reported outcome measures (PROMs) in palliative care—users’ cry for help. *Health Qual Life Outcomes* 2011;9:27.
21. Bausewein C, Daveson BA, Benalia A, Simon ST, Higginson IJ, on behalf of PRISMA. Outcome measurement in palliative care. The essentials. London: Cicely Saunders Institute, King’s College London, 2011.
22. Bausewein C, Le Grice C, Simon S, Higginson I. The use of two common palliative outcome measures in clinical care and research: a systematic review of POS and STAS. *Palliat Med* 2011;25: 304–313.
23. Higginson IJ, Evans CJ. What is the evidence that palliative care teams improve outcomes for cancer patients and their families? *Cancer J* 2010;16: 423–435.
24. Gysels MH, Evans C, Higginson IJ. Patient, caregiver, health professional and researcher views and experiences of participating in research at the end of life: a critical interpretive synthesis of the literature. *BMC Med Res Methodol* 2012;12:123.
25. Evans CJ, Harding R, Higginson IJ, on behalf of MORECare. ‘Best practice’ in developing and evaluating palliative and end-of-life care services: a meta-synthesis of research methods for the MORECare project. *Palliat Med* 2013 Jun 13. [Epub ahead of print].
26. Gysels M, Evans CJ, Lewis P, et al. MORECare research methods guidance development. Recommendations for ethical issues in palliative and end-of-life care research. *Palliat Med*, in press.
27. Preston NJ, Fayers P, Walters SJ, et al. Recommendations for managing missing data, attrition and response shift in palliative and end of life care research: part of the MORECare research methods guidance on statistical issues. *Palliat Med*, in press.
28. Higginson IJ, Evans CJ, Grande G, et al. Evaluating complex interventions in End of Life Care: the MORECare Statement on good practice generated by a synthesis of transparent expert consultations and systematic reviews *BMC Med*, in press.
29. Yardley L, Beyer N, Hauer K, et al. Recommendations for promoting the engagement of older people in activities to prevent falls. *Qual Saf Health Care* 2007;16:230–234.
30. Mularski RA, Dy SM, Shugarman LR, et al. A systematic review of measures of end-of-life care and its outcomes. *Health Serv Res* 2007;42: 1848–1870.
31. Fink A, Kosecoff J, Chassin M, Brook R. Consensus methods: characteristics and guidelines for use. *Am J Public Health* 1984;74:979–983.
32. Coffey A, Atkinson P. Making sense of qualitative data: Complementary research strategies. Thousand Oaks, CA: Sage Publications, 1996.
33. Murphy MK, Black NA, Lamping DL, et al. Consensus development methods, and their use in clinical guideline development. *Health Technol Assess* 1998;2:1–88.
34. Jones J, Hunter D. Consensus methods for medical and health services research. In: Mays N, Pope C, eds. *Qualitative research in health care*, 2nd ed. London: BMJ Publishing Group, 1999: 40–49.
35. Pope C, Ziebland S, Mays N. Analysing qualitative data. In: Pope C, Mays N, eds. *Qualitative research in health care*, 2nd ed. London: BMJ Books, 1999:75–101.
36. Garcia J, Evans J, Reshaw M. “Is there anything else you would like to tell us”—methodological issues in the use of free-text comments from postal surveys. *Qual Quant* 2004;38:113–115.

37. Lorenz KA, Lynn J, Dy SM, et al. Evidence for improving palliative care at the end of life: a systematic review. *Ann Intern Med* 2008;148:147–159.
38. Grande GE, Todd CJ. Why are trials in palliative care so difficult? *Palliat Med* 2000;14:69–74.
39. Rinck GC, van den Bos GA, Kleijnen J, et al. Methodologic issues in effectiveness research on palliative cancer care: a systematic review. *J Clin Oncol* 1997;15:1697–1707.
40. Daveson BA, Simon ST, Benalia H, et al. Are we heading in the same direction? European and African doctors' and nurses' views and experiences regarding outcome measurement in palliative care. *Palliat Med* 2012;26:242–249.
41. Higginson IJ, Simon ST, Benalia H, et al. Which questions of two commonly used multidimensional palliative care patient reported outcome measures are most useful? Results from the European and African PRISMA survey. *BMJ Support Palliat Care* 2012;2:36–42.
42. Mularski RA, Rosenfeld K, Coons SJ, et al. Measuring outcomes in randomized prospective trials in palliative care. *J Pain Symptom Manage* 2007;34:S7–S19.
43. Hearn J, Higginson IJ. Development and validation of a core outcome measure for palliative care: the palliative care outcome scale. Palliative Care Core Audit Project Advisory Group. *Qual Health Care* 1999;8:219–227.
44. Bruera E, Kuehn N, Miller MJ, Selmsler P, Macmillan K. The Edmonton Symptom Assessment System (ESAS): a simple method for the assessment of palliative care patients. *J Palliat Care* 1991;7:6–9.
45. Portenoy RK, Thaler HT, Kornblith AB, et al. The Memorial Symptom Assessment Scale: an instrument for the evaluation of symptom prevalence, characteristics and distress. *Eur J Cancer* 1994;30A:1326–1336.
46. Harding R, Selman L, Agupio G, et al. Validation of a core outcome measure for palliative care in Africa: the APCA African Palliative Outcome Scale. *Health Qual Life Outcomes* 2010;8:10.
47. Bausewein C, Fegg M, Radbruch L, et al. Validation and clinical application of the German version of the palliative care outcome scale. *J Pain Symptom Manage* 2005;30:51–62.
48. Van Soest-Poortvliet MC, van der Steen JT, Zimmerman S, et al. Psychometric properties of instruments to measure the quality of end-of-life care and dying for long-term care residents with dementia. *Qual Life Res* 2012;21:671–684.
49. Brandt HE, Deliens L, van der Steen JT, et al. The last days of life of nursing home patients with and without dementia assessed with the palliative care outcome scale. *Palliat Med* 2005;19:334–342.
50. Parker D, Hodgkinson B. A comparison of palliative care outcome measures used to assess the quality of palliative care provided in long-term care facilities: a systematic review. *Palliat Med* 2011;25:5–20.
51. Kutner JS, Bryant LL, Beaty BL, Fairclough DL. Symptom distress and quality-of-life assessment at the end of life: the role of proxy response. *J Pain Symptom Manage* 2006;32:300–310.
52. Hisamura K, Matsushima E, Nagai H, Mikami A. Comparison of patient and family assessments of quality of life of terminally ill cancer patients in Japan. *Psychooncology* 2011;20:953–960.
53. Jones JM, McPherson CJ, Zimmermann C, et al. Assessing agreement between terminally ill cancer patients' reports of their quality of life and family caregiver and palliative care physician proxy ratings. *J Pain Symptom Manage* 2011;42:354–365.
54. Tang ST, McCorkle R. Use of family proxies in quality of life research for cancer patients at the end of life: a literature review. *Cancer Invest* 2002;20:1086–1104.
55. Ingleton C, Faulkner A. Quality assurance in palliative care: some of the problems. *Eur J Cancer Care* 1995;4:38–44.
56. Groenvold M. Methodological issues in the assessment of health-related quality of life in palliative care trials. *Acta Anaesthesiol Scand* 1999;43:948–953.
57. Osoba D. Rationale for the timing of health-related quality-of-life (HQL) assessments in oncological palliative therapy. *Cancer Treat Rev* 1996;22(Suppl A):69–73.
58. Anderson JP, Kaplan RM, Schneiderman LJ. Effects of offering advance directives on quality adjusted life expectancy and psychological well-being among ill adults. *J Clin Epidemiol* 1994;47:761–772.
59. Bausewein C, Booth S, Gysels M, et al. Individual breathlessness trajectories do not match summary trajectories in advanced cancer and chronic obstructive pulmonary disease: results from a longitudinal study. *Palliat Med* 2010;24:777–786.