

## Original Article

**Cite this article:** Roydhouse J, Gutman R, Teno JM (2024) The potential impact of proxy reports for symptom experience and care quality and experience in advanced cancer. *Palliative and Supportive Care* 22(5), 904–912. <https://doi.org/10.1017/S1478951522001729>

Received: 04 June 2022  
Revised: 01 September 2022  
Accepted: 07 December 2022


**Keywords:**

Cancer; Experience; Quality; Proxy; Patient

**Author for correspondence:**

Jessica Roydhouse, Menzies Institute for Medical Research, University of Tasmania, Private Bag 23, Hobart, TAS 7001 Australia. Email: [jessica.roydhouse@utas.edu.au](mailto:jessica.roydhouse@utas.edu.au)

# The potential impact of proxy reports for symptom experience and care quality and experience in advanced cancer

Jessica Roydhouse, PH.D.<sup>1,2</sup> , Roeel Gutman, PH.D.<sup>3</sup> and Joan M. Teno, M.D., M.S.<sup>2,4</sup>

<sup>1</sup>Menzies Institute for Medical Research, University of Tasmania, Hobart, TAS, Australia; <sup>2</sup>Department of Health Services, Policy and Practice, Brown University School of Public Health, Providence, RI, USA; <sup>3</sup>Department of Biostatistics, Brown University School of Public Health, Providence, RI, USA and <sup>4</sup>Department of Medicine, Oregon Health and Science University, Portland, OR, USA

**Abstract**

**Objectives.** As the US tests models of care for the seriously ill, patient perceptions of the quality of care are important. Proxies are often needed for this group. We sought to understand the potential impact of proxy reports for the assessment of care quality and experience in cancer.

**Methods.** Secondary data analysis of a deidentified prospective study that included surveys of perceived care quality, including symptom management, from patients with advanced cancer receiving chemotherapy and their caregivers. Surveys were administered at diagnosis (time 1) and treatment (time 2), with top-box scoring used for analysis. Overall concordance was assessed using metrics including Gwet's AC1. The proportion of the highest scores by respondent type within 2 subgroups were examined: (1) symptom burden and (2) practice setting.

**Results.** Data from 83 dyads were analyzed. Proxies and patients frequently reported the highest scores for quality (time 1: proxies: 77% and patients: 80%). At time 1, 14% of proxies and 10% of patients reported an unmet need for symptom palliation. Most patients reporting an unmet need gave the top score for quality (75%), but fewer proxies did so (45%). Proxy and patient reports were similar within practice settings. Concordance was at least moderate (nearly all outcomes >0.5 and some >0.8) by Gwet's AC1.

**Significance of results.** Findings of at least moderate concordance and similar experience outcomes within subgroups suggest the use of proxies may not change estimates substantially. However, consideration should be taken when evaluating symptom management, particularly if such evaluations inform assessment of provider performance.

**Introduction**

Patient experience is increasingly an important part of the assessment of the quality of oncology care. This includes the evaluation of patients' perception of their experience through the Consumer Assessment of Healthcare Providers and Systems (CAHPS) Cancer Care Survey (Evensen et al. 2019) and interest in patient-reported outcome (PRO)-based performance measures (PRO-PMs) (Stover et al. 2020). A major challenge is that patients may be unable to self-report, and thus a proxy may be required. Previous research has demonstrated reasonable levels of concordance between paired proxy and patient reports for PROs (Roydhouse and Wilson 2017; Sneeuw et al. 2002; Tang and McCorkle 2002), although proxies reported worse patient health-related quality of life than patients (Roydhouse and Wilson 2017; Sneeuw et al. 2002; Tang and McCorkle 2002). Interestingly, paired concordance studies have shown that proxies report better care satisfaction (Castle 2005; Gasquet et al. 2003) and quality (Giovannetti et al. 2013) compared to patients.

Understanding the potential impact of proxy reports is important for fair comparisons across oncology practices. Previous research has demonstrated an association between poorer patient-perceived physician communication and unmet patient needs for symptom management (Walling et al. 2016). Furthermore, symptom burden is central to PRO-PMs (Stover et al. 2019a, 2020) and may inform comparisons of oncology practices (Stover et al. 2019a).

The objective of this study, which is a secondary data analysis, was to investigate the potential impact of proxy reports for the assessment of care quality and experience in cancer. The specific aims of this study were (1) to compare perceptions of care quality and experience and perceptions of symptom management and burden by respondent type (patient/proxy); (2) assess the concordance of patient and proxy reporting of symptom management and care quality and experience; (3) estimate the percentage of respondents giving the top score for each respondent type, within subgroups defined by symptom management/distress and for practice types. Earlier work using data from the original study using patient-reported problem scores suggested there

© The Author(s), 2023. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

was greater room for improvement in comprehensive care settings compared to private practice (Teno et al. 2009). As patient-reported experience can be used to differentiate between sites, we used practice type as a subgroup to examine the potential impact of proxy reports for such a comparison.

For experience outcomes, we focused on communication because of previous findings showing an association with symptom management (Walling et al. 2016), as well as perception of provider help with symptom management. For all outcomes, following CAHPS, we focused on the highest scores because experience outcomes are often skewed (Havyer et al. 2019), with most respondents selecting the top/highest values (Kemp et al. 2018; Takeshita et al. 2020). We hypothesized that scores for quality and experience would be high regardless of respondent type. Additionally, given interest in PRO-PMs for systemic therapy (Stover et al. 2020), we focused on the group of patients for whom both patients and proxies had received chemotherapy (either alone or with radiotherapy), as reported by both patients and proxies.

## Methods

### Data source and study design

This was a retrospective, secondary data analysis of an existing study. Because only de-identified data were used, the Brown University Institutional Review Board (IRB) deemed this study not human subjects research and IRB approval was not required. The original study, which was approved by the Brown University IRB, was a prospective cohort study that recruited patients diagnosed with advanced cancer in New Hampshire, Connecticut, and Rhode Island (Teno et al. 2009). Data were collected from 2006 to 2007 at 2 time points: after initial diagnosis or progression to advanced disease and after receiving treatment (Teno et al. 2009). There were 119 patients (of 206 recruited) (Teno et al. 2009) who had caregivers willing to complete questionnaires at time 1, with data from 99 dyads available at time 2. The study population consisted of 83 dyads where both members reported the patient receiving chemotherapy and had questionnaire data at both time points. We were interested in investigating these questions at more than one time point because studies evaluating proxy–patient concordance for PROs related to patient health have not had consistent results about concordance over time (Milne et al. 2006). As perceptions of health may change over time, it is plausible that perceptions of care experience may change as well, particularly because health status is an important predictor of satisfaction with care (Hall et al. 1993).

### Measures

In both surveys, proxies were told that they were being asked to answer questions about the patient's experiences with cancer care. No explicit directions were provided regarding perspective taking (e.g., use of their own perspective or answering as they believe the patient would have) (Pickard and Knight 2005).

### Quality and experience outcomes

Prior to commencing questions in this section of the questionnaire, patients were given the following instructions at time 1: "Now I would like you to think about your overall experience with your oncologist, your cancer care providers and oncology office staff by rating some aspects of the care you have received. For the next set of questions, I will be asking you to respond using a scale from 0 to 10, where 0 means the worst care possible and 10 means the best

care possible." Proxy instructions were the same, but with "your" replaced by "his/her," referring to the patient. The instructions were the same at time 2, except for both respondents, the first sentence ended with "within the past month."

Quality of care was assessed with a 5-point scale ranging from poor to excellent. Respondents were asked, "Overall, how would you rate the care that you [PATIENT] have received from your (his/her) oncologist." The same question was administered at time 1 and time 2. Communication was assessed by the question, "Using a scale of 0–10, how well do your [PATIENT] cancer care providers communicate with you [PATIENT] and your (his/her) family/friends about your (his/her) cancer and the likely outcomes of care?"; for patients and proxies, 0 was the worst score and 10 was the best score. The same question was administered at time 1 and time 2. Experience with provider help for symptom control was examined with the question, "Using a scale of 0–10, how well do your [PATIENT] cancer care providers make sure that your (his/her) symptoms (e.g., pain and shortness of breath) are controlled to a degree that is acceptable to you (him/her)?" At time 2, the question was the same except "nausea" was added as another example of a symptom for both respondents: "(e.g., pain, shortness of breath, and nausea)."

### Assessment of symptoms

Pain, trouble breathing, and feelings of anxiety or distress were assessed at both time points. Diarrhea and nausea/vomiting were assessed at time 2 but not time 1. If respondents indicated a symptom was present in the past month (y/n), they answered follow-up questions including how much the symptom bothered or distressed the patient (not at all to very much) and the amount of help received for the symptom (less than needed/about the right amount/more than needed).

For this analysis, we operationalized symptom burden and management in 2 ways. First, using only symptoms that were measured at both time points (pain, trouble breathing, and anxiety/distress), we created a 2-level variable to determine the presence and management of symptoms (at least 1 symptom with an undesired level of help/no symptoms or a desired level of help if symptoms present). This derived variable measures unmet needs for symptom management. Second, because pain is one of the symptoms that met inclusion criteria for PRO-PMs (Stover et al. 2019b) and the literature has examples of generally accurate (Milne et al. 2006) and less accurate (Montgomery et al. 2021) proxy reporting of symptom distress, we created a binary variable to indicate whether patients had any level of distress or bother from pain (no pain or pain without distress or bother/pain and some level of distress or bother). A top score for this variable indicated no pain or pain without distress/bother. This derived variable measures symptom burden.

### Variables describing study sample

Except for cancer and practice type (comprehensive cancer center or private practice), all other variables were patient- or proxy-reported. Detailed information about the proxy–patient relationship was also elicited from the proxy with questions about the frequency of assisting the patient, discussion of medical condition and attendance at oncologist visits (never to always), and extent of involvement in cancer treatment decisions (not at all to very much). Proxies also rated their confidence in understanding the impact of cancer on the patient's health and of how treatment decisions were made on 0–10 scales (not to very). Proxies and patients also rated their own overall health status (poor to excellent), level of distress (0–10), and level of psychological distress and well-being using the

validated Mental Health Inventory-5 (MHI-5) (Berwick et al. 1991; Rumpf et al. 2001). For consistency, all health-related questions and scales were transformed so that higher scores indicated better responses.

### Data analyses

Frequencies, means, and standard deviations were used to describe the data. Because of the small sample size, we focused on descriptive rather than inferential statistics for analysis. The description of perception of quality and care experience and symptom management and burden (Aim 1) was undertaken using descriptive statistics. Kendall's tau was used to evaluate the association between proxy and patient responses on the rating scales for provider communication and efforts with symptom control. Concordance (Aim 2) was assessed using several metrics. Although kappa is widely used to evaluate inter-rater concordance (Shankar and Bangdiwala 2014), its performance can be affected by prevalence (Byrt et al. 1993; Feinstein and Cicchetti 1990; Flight and Julious 2015), which is of concern given that experience data are skewed. Our primary measure was therefore Gwet's AC1 (Gwet 2008), but for comprehensiveness, we assessed Cohen's kappa (Cohen 1960), percent agreement,  $p_{\text{pos}}$  and  $p_{\text{neg}}$  (Cicchetti and Feinstein 1990). Kappa, Gwet's AC1, and percent agreement were calculated using the irrCAC package (Gwet 2019).  $p_{\text{pos}}$  was calculated as the proportion of agreement for top (best) scores and  $p_{\text{neg}}$  for <top scores, using Cicchetti and Feinstein's formulas (Cicchetti and Feinstein 1990).

For the subgroup analyses (Aim 3), we estimated the percentage of top scores by respondent type for each level of symptom management and burden at each time point. We repeated this analysis for subgroups defined by practice type. For all aims, each outcome was evaluated separately. R Studio (v.4.0.1) was used for all analyses.

### Missing data

Missing values were infrequent, with only 1–3 outcome scores missing at each time point. Available case analysis was used.

## Results

### Study population

Most patients received care from a private practice (67%) rather than a comprehensive care site (33%). Patients were older than their caregivers (mean age 65.6 [SD = 10.7] compared to 59.4 [SD = 12.8]), whereas a higher proportion of proxies had received at least a university-level education compared to patients (28% vs. 20%). Most proxies were the patient's partner or spouse (69%), and nearly all reported being in contact with the patient every day (95%). Proxies reported being highly engaged in care, with the majority "always" discussing the patient's medical condition with them (54%) and attending oncologist visits (63%), as well as being "very much" involved in patient treatment decisions (76%) (Table 1).

### Symptom management (receipt of desired level of help for symptoms)

At both times 1 and 2, a higher percentage of proxies than patients reported inappropriate help for patient symptoms (time 1: 14% vs. 10% and time 2: 12% vs. 6%) (Table 1). Concordance as measured by Gwet's AC (Gwet 2008) for this outcome was high

(0.82 at time 1 and 0.86 at time 2); similarly, percent agreement and  $p_{\text{pos}}$  were  $>0.80$  at both time points.  $p_{\text{neg}}$  was  $<0.5$  at both time points. Cohen's kappa (Cohen 1960) was low at both time points (0.32 at time 1 and 0.22 at time 2) (Table 2).

### Presence of pain distress

Higher percentages of proxies than patients reported that patients experienced some distress from pain, at both time 1 (59% vs. 52%) and time 2 (59% vs. 57%). However, at both time points, the majority of each respondent type reported distress from pain (Table 1). Concordance as measured by Gwet's AC was higher at time 2 (0.86) compared to time 1 (0.67), though  $>0.5$  at both time points.  $p_{\text{pos}}$  was  $>0.80$  at both time points (time 1: 0.81 and time 2: 0.91), as was percent agreement.  $p_{\text{neg}}$  (time 1: 0.85 and time 2: 0.94) and Cohen's kappa were likewise high at both time points (time 1: 0.66 and time 2: 0.85) (Table 2).

### Care quality

Nearly all proxies and patients gave the top scores for quality at both time points (t1: 77% vs. 80% and t2: 76% vs. 78%) (Table 1). Gwet's AC showed high concordance (time 1: 0.67 and time 2: 0.70), as did  $p_{\text{pos}}$  (time 1: 0.86 and time 2: 0.87); percent agreement was  $>0.7$  at both time points. However,  $p_{\text{neg}}$  (time 1: 0.47 and time 2: 0.56) and Kappa were lower (time 1: 0.33 and time 2: 0.43) (Table 2).

### Provider communication

At both time points, proxy and patient scores were consistently high, with mean scores of  $\sim 9$  regardless of respondent type at both time points (Table 1). Interestingly, although most respondents gave very high scores as indicated by the mean scores, at both time points, the score range reported by patients was wider (time 1: 0–10 vs. 5–10 and time 2: 0–10 vs. 4–10) (data not shown). Furthermore, lower percentages of proxies gave the top scores for communication at time 1 than those of patients (58% vs. 69%), but the reverse was true at time 2 (53% vs. 49%) (Table 1). Correlation for the ordinal provider communication outcomes between respondents was similar at both time points: t1:  $\tau = 0.366$  and t2:  $\tau = 0.345$  (data not shown). Concordance as measured by Gwet's AC was  $<0.5$  for this outcome at both time points (time 1: 0.44 and time 2: 0.26), although  $p_{\text{pos}}$  was  $>0.5$  (time 1: 0.76 and time 2: 0.64), as was  $p_{\text{neg}}$  (time 1: 0.59 and time 2: 0.62). Kappa was likewise low (time 1: 0.36 and time 2: 0.26) (Table 2), while percent agreement was  $>0.6$  at both time points.

### Provider help with symptom control

Both patients and proxies gave very high scores for provider help with symptom control at both time points, with mean scores of  $\sim 9$  reported at every time point regardless of respondent type (Table 1). At both times 1 and 2, over 50% of respondents regardless of type gave the highest score for provider efforts with symptom control (t1: 64% vs. 67%; t2: 59% vs. 59%) (Table 1). Correlation for providers' efforts with symptom control outcome increased over time (t1:  $\tau = 0.215$ , t2:  $\tau = 0.486$ ) (data not shown). Concordance as measured by Gwet's AC likewise showed an increase over time (time 1: 0.38 and time 2: 0.54), as did Cohen's kappa (time 1: 0.18 and time 2: 0.51).  $p_{\text{pos}}$  was high at both time points (time 1: 0.74 and time 2: 0.8), and  $p_{\text{neg}}$  increased over time (time 1: 0.44 and time 2: 0.71), as did percent agreement (time 1: 0.65 and time 2: 0.76).

**Table 1.** Study sample characteristics and outcomes (N = 83)

Characteristic	Patients	Proxies	Patients	Proxies
<i>Sociodemographic/clinical characteristics</i>	Time 1	Time 1	Time 2	Time 2
Age in years, mean (SD)	65.6 (10.7)	59.4 (12.8)		
Education				
1: <HS	11 (13%)	7 (8%)		
2: <Uni	55 (66%)	53 (64%)		
3: Uni+	17 (20%)	23 (28%)		
Practice type				
0: Comprehensive care site	27 (33%)			
1: Private practice	56 (67%)			
Marital status				
1: Married	57 (69%)			
2: Not married	24 (29%)			
Missing	2 (2%)			
Gender				
1: Male	41 (49%)			
2: Female	42 (51%)			
Cancer diagnosis				
1: Lung	30 (36%)			
2: Colorectal	21 (25%)			
3: Breast	13 (16%)			
4: Other	19 (23%)			
<i>Health characteristics</i>				
Mental health: MHI-5 (0–100, higher = better), mean (SD)	T1: 74.8 (19.1)	T1: 68.2 (19.0)	T2: 75.2 (17.8) (3 missing)	T2: 71.6 (18.1) (1 missing)
Distress in past week (0–10, higher = better)	T1: 7.1 (2.8)	T1: 6.1 (3.0)	T2: 7.1 (2.8) (2 missing)	T2: 6.3 (2.8)
Overall health rating				
1: Poor	5 (6%)	0 (0%)	6 (7%)	0 (0%)
2: Fair	11 (13%)	10 (12%)	19 (23%)	8 (10%)
3: Good	33 (40%)	29 (35%)	28 (34%)	34 (41%)
4: Very Good	27 (33%)	36 (43%)	23 (28%)	34 (41%)
5: Excellent	7 (8%)	8 (10%)	5 (6%)	7 (8%)
Missing			2 (2%)	
<i>Proxy characteristics</i>				
Proxy/patient relationship				
1: Partner/spouse		57 (69%)		
2: Child		18 (22%)		
3: Other		8 (10%)		
Frequency of proxy/patient contact				
1: 1–3 d/week		1 (1%)		
2: 4–6 d/week		2 (2%)		
3: 7 d/week		79 (95%)		
Missing		1 (1%)		
Frequency of assisting patient with transport or shopping				

(Continued)

**Table 1.** (Continued.)

Characteristic	Patients	Proxies	Patients	Proxies
1: Never		9 (11%)		
2: Sometimes		17 (20%)		
3: Usually		16 (19%)		
4: Always		39 (47%)		
Missing		2 (2%)		
Frequency of discussion of patient's medical condition				
1: Never		1 (1%)		
2: Sometimes		16 (19%)		
3: Usually		21 (25%)		
4: Always		45 (54%)		
Frequency of proxy attendance at oncologist visits				
1: Never		0 (0%)		
2: Sometimes		11 (13%)		
3: Usually		20 (24%)		
4: Always		52 (63%)		
Proxy involvement in patient treatment decisions				
1: Not at all		3 (4%)		
2: A little bit		5 (6%)		
3: Somewhat		12 (14%)		
4: Very much		63 (76%)		
Proxy is the one if patient would count on for help if confined to bed		81 (98%)		
Proxy confidence in understanding patient's cancer and its impact on health (0–10 scale, higher = better), mean (SD)		8.9 (1.5)		
Proxy confidence in understanding how treatment decisions are made (0–10 scale, higher = better), mean (SD)		9.2 (1.5) (1 missing)		
How well proxy knows patient (0–10 scale, higher = better), mean (SD)		9.8 (0.6)		
<i>Outcomes</i>				
Symptom				
Correct help level for symptoms/no symptoms	75 (90%)	71 (86%)	77 (93%)	73 (88%)
Any symptom with wrong help level	8 (10%)	12 (14%)	5 (6%)	10 (12%)
Missing			1 (1%)	
No pain, or if pain then no distress	40 (48%)	34 (41%)	35 (42%)	34 (41%)
Pain and some level of distress	43 (52%)	49 (59%)	47 (57%)	49 (59%)
Missing			1 (1%)	
<i>Care quality and experience</i>				
Quality of care: top score (excellent)	66 (80%)	64 (77%)	65 (78%)	63 (76%)
Missing		1 (1%)	1 (1%)	1 (1%)
Rating for communication (0–10, worst to best)	9.1 (1.8)	9.2 (1.1)	9.0 (1.6) (1 missing)	8.9 (1.5) (1 missing)
Top score for provider communication	57 (69%)	48 (58%)	41 (49%)	44 (53%)
Missing			1 (1%)	1 (1%)
Rating for provider help with symptom control (0–10, worst to best)	9.2 (1.5) (3 missing)	9.3 (1.2) (1 missing)	9.4 (0.9) (2 missing)	9.3 (1.2) (1 missing)
Top score for provider help: symptom control	56 (67%)	53 (64%)	49 (59%)	49 (59%)
Missing	3 (4%)	1 (1%)	2 (2%)	1 (1%)



**Table 2.** Concordance (binary scores)

Time/aspect	Kappa	Gwet's AC	Percent agreement	$p_{pos}$	$p_{neg}$
<b>Time 1</b>					
Symptom management	0.32	0.82	0.86	0.92	0.4
Pain distress	0.66	0.67	0.83	0.81	0.85
Quality of care	0.33	0.67	0.78	0.86	0.47
Communication	0.36	0.44	0.70	0.76	0.59
Provider help with symptom control	0.18	0.38	0.65	0.74	0.44
<b>Time 2</b>					
Symptom management	0.22	0.86	0.88	0.93	0.29
Pain distress	0.85	0.86	0.93	0.91	0.94
Quality of care	0.43	0.70	0.80	0.87	0.56
Communication	0.26	0.26	0.63	0.64	0.62
Provider help with symptom control	0.51	0.54	0.76	0.8	0.71

**Reports of quality and experience within subgroups**

*Appropriateness of help received for symptoms*

Most patients who reported receiving the wrong level of help for a symptom nonetheless reported the highest quality score ( $t1$ : 75% and  $t2$ : 80%), whereas this was not the case for proxies ( $t1$ : 45% and  $t2$ : 40%). However, this pattern was not seen for the communication scores. Among the subgroup of patients or proxies reporting inappropriate help for symptoms, a higher percentage of proxies than patients reported the highest communication score at time 1 (33% vs. 13%) but not at time 2 (30% vs. 40%). For the provider help with symptom control outcome, a higher percentage of patients than proxies endorsed a top score even when inappropriate help for symptoms was received, both at time 1 (38% vs. 25%) and time 2 (60% vs. 30%) (Table 3).

*Pain distress*

Regardless of the presence or absence of pain distress, >50% proxies and patients reported the highest scores for all quality and experience outcomes at time 1. At time 2, >50% of both respondent types continued to report the highest possible quality and experience scores across distress levels, except for the communication score, where <50% of proxies (48%) and patients (49%) who reported no distress gave the highest score (Table 3).

*Practice type*

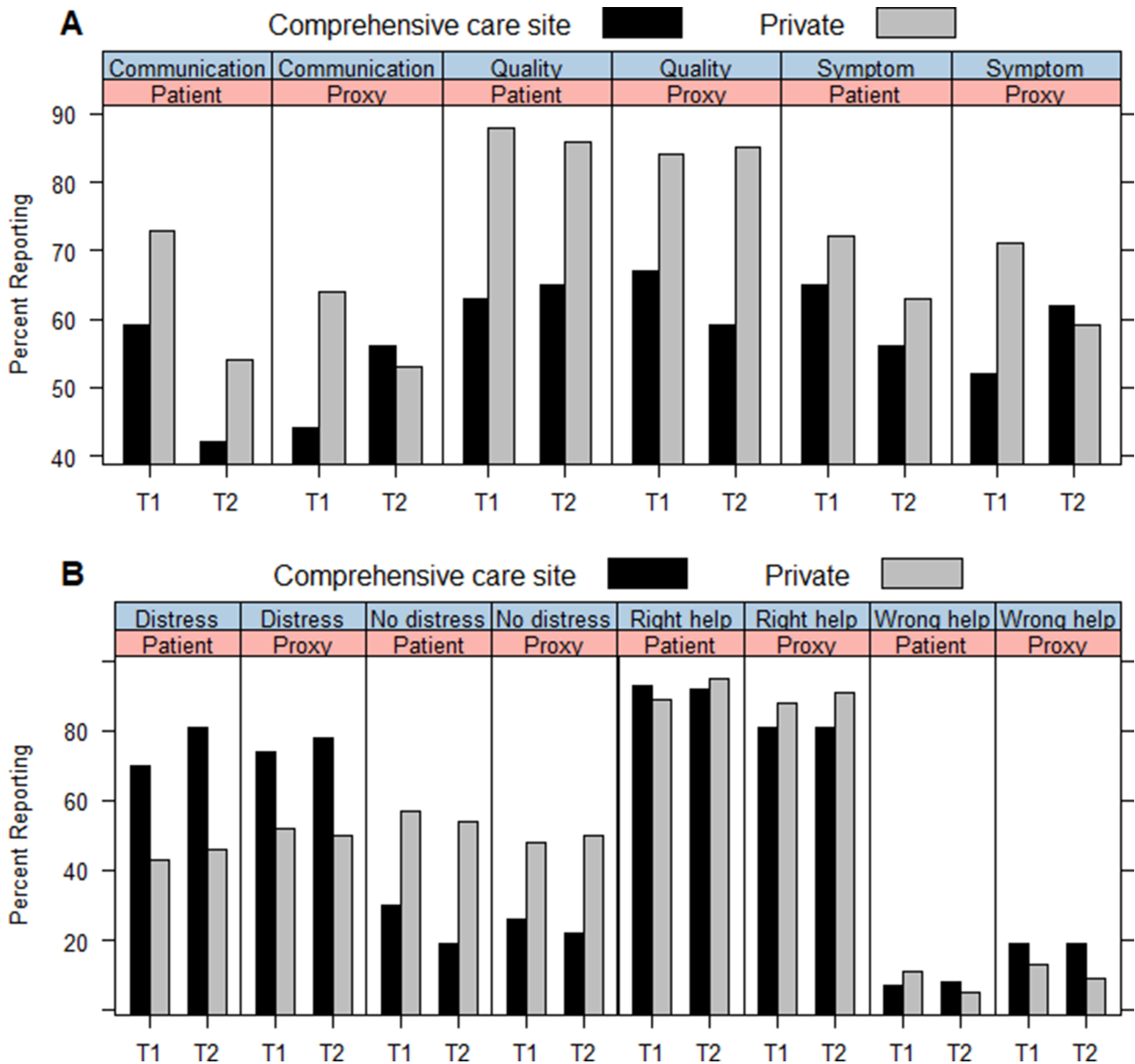
The percentage of top scores reported for the quality outcome was similar for patients and proxies, regardless of practice type, with higher percentages reported for private practice at both time points (Figure 1A). However, there were differences between respondent types for the symptom and communication outcomes. For both outcomes, the percentage of patients giving the highest score for comprehensive care sites was higher at time 1 than time 2, whereas the reverse was true for proxies. These between-respondent differences were not seen for private practice.

Mean scores for the communication and symptom outcomes were high across respondent and practice types. At time 1, mean

**Table 3.** Care quality and experience within strata of symptom experience, by respondent type

Symptom experience	Top quality score (%)	Top communication score (%)	Top symptom score (%)
<i>Symptom management</i>			
Time 1 – patients			
Appropriate help or no symptoms	60 (80%)	56 (75%)	53 (74%)
Inappropriate help	6 (75%)	1 (13%)	3 (38%)
Time 1 – proxies			
Appropriate help or no symptoms	59 (83%)	44 (62%)	50 (71%)
Inappropriate help	5 (45%)	4 (33%)	3 (25%)
Time 2 – patients			
Appropriate help or no symptoms	61 (79%)	39 (51%)	46 (61%)
Inappropriate help	4 (80%)	2 (40%)	3 (60%)
Time 2 – proxies			
Appropriate help or no symptoms	59 (82%)	41 (57%)	46 (64%)
Inappropriate help	4 (40%)	3 (30%)	3 (30%)
<i>Pain distress</i>			
Time 1 – patients			
No or unknown pain or pain distress	31 (78%)	30 (75%)	28 (76%)
Pain causing some level of distress	35 (81%)	27 (63%)	28 (65%)
Time 1 – proxies			
No or unknown pain or pain distress	27 (79%)	19 (56%)	22 (67%)
Pain causing some level of distress	37 (77%)	29 (59%)	31 (63%)
Time 2 – patients			
No or unknown pain or pain distress	29 (83%)	17 (49%)	21 (60%)
Pain causing some level of distress	36 (77%)	24 (51%)	28 (61%)
Time 2 – proxies			
No or unknown pain or pain distress	27 (82%)	16 (48%)	20 (61%)
Pain causing some level of distress	36 (73%)	28 (57%)	29 (59%)

communication scores for proxies and patients were only slightly higher for private practice (proxies: mean = 9.4, SD = 1.1 and patients: mean = 9.2, SD = 2.0) compared to comprehensive care sites (proxies: mean = 9.0, SD = 1.1 and patients: mean = 9.0, SD = 1.4). Similarly, respondent scores were close at time 2 (data not shown). The mean scores for symptom outcomes at time 1 were high for both proxies and patients for private practice (mean = 9.5, SD = 1.0 and mean = 9.3, SD = 1.4, respectively) and for comprehensive care sites (mean = 9.0, SD = 1.6 for both). Similar results were seen for time 2 (data not shown).



**Fig. 1.** Comparison of quality, experience, and symptoms across practice types. (A) Comparison of quality and care experience outcomes. (B) Comparison of symptom burden and management outcomes.

Comparing private practices and comprehensive care sites for symptom management and burden yielded similar results, regardless of respondent type or time point (Figure 1B). Specifically, higher percentages of reports of pain distress were seen for comprehensive care sites, and proxy and patient reports of these were similar. High percentages of both proxies and patients reported “correct level of help for symptoms” for comprehensive care sites and private practices. However, at time 1, there was a higher percentage of patients reporting an inappropriate level of help in private practices compared to comprehensive care sites, but the reverse was true at time 2. In contrast, for proxies, the percentage was consistently higher for comprehensive care sites. Furthermore, a higher percentage of proxies reported receipt of inappropriate help for symptoms compared to patients regardless of site, although in general, this was not frequent (<20% for all time points and respondent types).

**Discussion**

Most participants, regardless of respondent type, gave the highest possible quality and experience scores at both time points. At the same time, at least 10% (depending on respondent type) of participants indicated an unmet need for symptom management. Concordance was generally moderate to good. In subgroup analyses, there were some differences between proxies and patients in terms of reporting the highest score, but these were not consistent. Such differences have the potential to be influential in studies where the proportions of proxies differ between arms or sites.

Of note is that concordance for perceived communication was <0.5 despite proxies and patients giving generally high scores at both time points. Patient-proxy concordance on communication items was lower compared to concordance on more objective items in another study (Giovannetti et al. 2013); however, as quality is

not necessarily more objective than communication, it is unclear why communication would have lower concordance in our study. Because the communication question investigates the “likely outcomes of cancer care,” differences in expectations or perception of likely outcomes of cancer care between patients and proxies may be a possible explanation.

Previous research involving patient-proxy dyads to compare symptom assessment has generally found that proxies tend to overestimate patient symptoms (Ma et al. 2021; McPherson et al. 2008; Silveira et al. 2010; Winters-Stone et al. 2014), although there are also studies with findings of high levels of patient-proxy agreement (Akin and Durna 2013; Armstrong et al. 2012) or of no differences (Miller et al. 2015; Yeager et al. 2022). In this study, we found that higher percentages of proxies reported pain distress and the receipt of inappropriate help for symptoms, but the percentages of proxies and patients reporting pain distress were closer at time 2 than time 1, whereas the reverse was true for inappropriate help for symptoms. Additional longitudinal studies are needed to understand how proxy reporting may change over time.

Growing interest in patient-reported symptom measures as part of performance measurement in oncology (Stover et al. 2020) highlights the potential importance of proxy ratings and potential differences between proxies and patients. The exclusion of patients who cannot report their own symptoms may result in a biased population, but lack of adjustment for proxy-patient differences may disadvantage services with higher proportions of proxies in between-service comparisons. In some surveys, the frequency of proxy use can vary across racial/ethnic groups (Pinheiro et al. 2015); if the sociodemographic characteristics of the population served vary by services or sites, this could affect the frequency of proxy use and the measured experience for those services or sites.

Additionally, patient experience measures can be used to inform improvements at a service level (Manalili and Santana 2021). Our findings, like those of Havyer et al. (2019), suggest agreement is not as high for scores other than the highest scores, which can pose challenges for the use of the data to inform quality improvement efforts (Havyer et al. 2019) in the absence of guidelines for approaches that adjust for proxy reports. The development of such guidelines is an important area for future research. In addition to analytic methods, other considerations for data collection and future work include factors that may affect proxy reporting. Previous research has indicated that the proxy-patient relationship and the proxy's engagement in care are associated with proxy reports of care experience and quality (Roydhouse et al. 2018a).

This study has several limitations. The sample size was relatively small and drawn from practices in a few north-eastern US states, which may limit generalizability. Additionally, the data were collected several years ago, and cancer treatment has changed in recent years. However, chemotherapy remains a mainstay of treatment, and the methodological focus of the study makes the age of the data less concerning. Additional limitations pertaining to the age of the data include improvements in symptom management and greater focus on patient-centered care and communication.

Additionally, in this study, proxies were instructed to report about the patient's experiences with care, but it is not clear what perspective they may have taken when doing so. Furthermore, this study has several strengths, including the assessment at more than one time point. The relative paucity of longitudinal data on proxies has been recognized before (Roydhouse and Wilson 2017; Sneeuw et al. 2002), and because there is an increasing interest in PRO assessment throughout the cancer journey, studies that are not limited to single time points are important.

In conclusion, our findings suggest that the use of proxies to report on care quality and experience outcomes may not change estimates substantially. Proxy endorsement of top scores was only substantially lower for the subgroup of patients with perceived inappropriate symptom help. Should this group vary in size across sites, it is possible that estimates may be affected. Furthermore, although data from both proxies and patients were available in this study, this is unlikely to be the case in practice, as proxies will be required when patients are too ill to self-report. Because patients requiring proxies tend to be older and in poorer health (Roydhouse et al. 2018b), development of methods to analyze datasets with information from both patients and proxies is important, particularly if these assessments inform evaluations of provider performance or are used for quality improvement.

**Funding.** J.R. is supported by a Select Foundation Fellowship. This is a secondary data analysis, and the original study was supported by National Cancer Institute grant no. CA097358 R01.

**Conflicts of interest.** Jessica Roydhouse reports personal fees from Amgen, outside the submitted work, and consultancies with University of Birmingham Enterprise, outside the submitted work. Roeel Gutman served as an expert witness to Janssen/J&J on an unrelated matter.

## References

- Akin S and Durna Z (2013) A comparative descriptive study examining the perceptions of cancer patients, family caregivers, and nurses on patient symptom severity in Turkey. *European Journal of Oncology Nursing* 17(1), 30–37. doi:10.1016/j.ejon.2012.02.002
- Armstrong TS, Wefel JS, Gning I, et al. (2012) Congruence of primary brain tumor patient and caregiver symptom report. *Cancer* 118(20), 5026–5037. doi:10.1002/cncr.27483
- Berwick DM, Murphy JM, Goldman PA, et al. (1991) Performance of a five-item mental health screening test. *Medical Care* 29(2), 169–176. doi:10.1097/00005650-199102000-00008
- Byrt T, Bishop J and Carlin JB (1993) Bias, prevalence and kappa. *Journal of Clinical Epidemiology* 46(5), 423–429. doi:10.1016/0895-4356(93)90018-V
- Castle N (2005) Are family members suitable proxies for transitional care unit residents when collecting satisfaction information? *International Journal for Quality in Health Care* 17(5), 439–445. doi:10.1093/intqhc/mzi056
- Cicchetti DV and Feinstein AR (1990) High agreement but low kappa: II. Resolving the paradoxes. *Journal of Clinical Epidemiology* 43(6), 551–558. doi:10.1016/0895-4356(90)90159-M
- Cohen J (1960) A coefficient of agreement for nominal scales. *Educational and Psychological Measurement* 20(1), 37–46. doi:10.1177/00131644600200104
- Evensen CT, Yost KJ, Keller S, et al. (2019) Development and testing of the CAHPS cancer care survey. *Journal of Oncology Practice* 15(11), e969–e978. doi:10.1200/JOP.19.00039
- Feinstein AR and Cicchetti DV (1990) High agreement but low kappa: I. The problems of two paradoxes. *Journal of Clinical Epidemiology* 43(6), 543–549. doi:10.1016/0895-4356(90)90158-L
- Flight L and Julious SA (2015) The disagreeable behaviour of the kappa statistic. *Pharmaceutical Statistics* 14(1), 74–78. doi:10.1002/pst.1659
- Gasquet I, Dehe S, Gaudebout P, et al. (2003) Regular visitors are not good substitutes for assessment of elderly patient satisfaction with nursing home care and services. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences* 58(11), 1036–1041. doi:10.1093/gerona/58.11.M1036
- Giovannetti ER, Reider L, Wolff JL, et al. (2013) Do older patients and their family caregivers agree about the quality of chronic illness care? *International Journal for Quality in Health Care* 25(5), 515–524. doi:10.1093/intqhc/mzt052
- Gwet KL (2008) Computing inter-rater reliability and its variance in the presence of high agreement. *British Journal of Mathematical and Statistical Psychology* 61(Pt 1), 29–48. doi:10.1348/000711006X126600



- Gwet KL (2019) Package “irrCAC”. <https://cran.r-project.org/web/packages/irrCAC/index.html> (accessed 31 January 2022).
- Hall JA, Milburn MA and Epstein AM (1993) A causal model of health status and satisfaction with medical care. *Medical Care* **31**(1), 84–94. doi:10.1097/00005650-199301000-00007
- Havyer RD, van Ryn M, Wilson PM, et al. (2019) Concordance of patient and caregiver reports on the quality of colorectal cancer care. *Journal of Oncology Practice* **15**(11), e979–e988. doi:10.1200/JOP.19.00136
- Kemp KA, Ahmed S, Quan H, et al. (2018) Family experiences of pediatric inpatient care in Alberta, Canada: Results from the child HCAHPS survey. *Hospital Pediatrics* **8**(6), 338–344. doi:10.1542/hpeds.2017-0191
- Manalili K, Santana MJ and ISOQOL PROMs/PREMs in Clinical Practice Implementation Science Work Group (2021) Using implementation science to inform the integration of electronic patient-reported experience measures (ePREMs) into healthcare quality improvement: Description of a theory-based application in primary care. *Quality of Life Research* **30**(11), 3073–3084. doi:10.1007/s11136-020-02588-1
- Ma X, Yu W, Lu Y, et al. (2021) Congruence of cancer pain experience between patients and family caregivers and associated factors: A multicenter cross-sectional study in China. *Supportive Care in Cancer* **29**(10), 5983–5990. doi:10.1007/s00520-021-06156-2
- McPherson CJ, Wilson KG, Lobchuk MM, et al. (2008) Family caregivers’ assessment of symptoms in patients with advanced cancer: Concordance with patients and factors affecting accuracy. *Journal of Pain and Symptom Management* **35**(1), 70–82. doi:10.1016/j.jpainsymman.2007.02.038
- Miller LM, Lyons KS and Bennett JA (2015) Incongruent perceptions of pain and physical function among families living with lung cancer. *Supportive Care in Cancer* **23**(9), 2755–2762. doi:10.1007/s00520-015-2640-x
- Milne DJ, Mulder LL, Beelen HC, et al. (2006) Patients’ self-report and family caregivers’ perception of quality of life in patients with advanced cancer: How do they compare? *European Journal of Cancer Care* **15**(2), 125–132. doi:10.1111/j.1365-2354.2005.00639.x
- Montgomery KE, Vos K, Raybin JL, et al. (2021) Comparison of child self-report and parent proxy-report of symptoms: Results from a longitudinal symptom assessment study of children with advanced cancer. *Journal for Specialists in Pediatric Nursing* **26**(3), e12316. doi:10.1111/jspn.12316
- Pickard AS and Knight SJ (2005) Proxy evaluation of health-related quality of life: A conceptual framework for understanding multiple proxy perspectives. *Medical Care* **43**(5), 493–499. doi:10.1097/01.mlr.0000160419.27642.a8
- Pinheiro LC, Wheeler SB, Chen RC, et al. (2015) The effects of cancer and racial disparities in health-related quality of life among older Americans: A case-control, population-based study. *Cancer* **121**(8), 1312–1320. doi:10.1002/cncr.29205
- Roydhouse JK, Gutman R, Keating NL, et al. (2018a) The association of proxy care engagement with proxy reports of patient experience and quality of life. *Health Services Research* **53**(5), 3809–3824. doi:10.1111/1475-6773.12980
- Roydhouse JK, Gutman R, Keating NL, et al. (2018b) Differences between proxy and patient assessments of cancer care experiences and quality ratings. *Health Services Research* **53**(2), 919–943. doi:10.1111/1475-6773.12672
- Roydhouse JK and Wilson IB (2017) Systematic review of caregiver responses for patient health-related quality of life in adult cancer care. *Quality of Life Research* **26**(8), 1925–1954. doi:10.1007/s11136-017-1540-6
- Rumpf H-J, Meyer C, Hapke U, et al. (2001) Screening for mental health: Validity of the MHI-5 using DSM-IV Axis I psychiatric disorders as gold standard. *Psychiatry Research* **105**(3), 243–253. doi:10.1016/S0165-1781(01)00329-8
- Shankar V and Bangdiwala SI (2014) Observer agreement paradoxes in 2x2 tables: Comparison of agreement measures. *BMC Medical Research Methodology* **14**, 100. doi:10.1186/1471-2288-14-100
- Silveira MJ, Given CW, Given B, et al. (2010) Patient-caregiver concordance in symptom assessment and improvement in outcomes for patients undergoing cancer chemotherapy. *Chronic Illness* **6**(1), 46–56. doi:10.1177/1742395309359208
- Sneeuw KC, Sprangers MA and Aaronson NK (2002) The role of health care providers and significant others in evaluating the quality of life of patients with chronic disease. *Journal of Clinical Epidemiology* **55**(11), 1130–1143. doi:10.1016/S0895-4356(02)00479-1
- Stover A, Urick B, Deal A, et al. (2019a) ON5 development and testing of patient-reported outcome performance measures (PRO-PMs) for comparing oncology practices. *Value in Health* **22**, S413. doi:10.1016/j.jval.2019.09.085
- Stover AM, Urick B, Deal AM, et al. (2019b) Development and testing of patient-reported outcome performance measures (PRO-PMs) for oncology practice. *Journal of Clinical Oncology* **37**, 173. doi:10.1200/JCO.2019.37.27\_suppl.173
- Stover AM, Urick BY, Deal AM, et al. (2020) Performance measures based on how adults with cancer feel and function: Stakeholder recommendations and feasibility testing in six cancer centers. *JCO Oncology Practice* **16**(3), e234–e250. doi:10.1200/JOP.19.00784
- Takeshita J, Wang S, Loren AW, et al. (2020) Association of racial/ethnic and gender concordance between patients and physicians with patient experience ratings. *JAMA Network Open* **3**(11), e2024583. doi:10.1001/jamanetworkopen.2020.24583
- Tang ST and McCorkle R (2002) Use of family proxies in quality of life research for cancer patients at the end of life: A literature review. *Cancer Investigation* **20**, 1086–1104. doi:10.1081/CNV-120005928
- Teno JM, Lima JC and Lyons KD (2009) Cancer patient assessment and reports of excellence: Reliability and validity of advanced cancer patient perceptions of the quality of care. *Journal of Clinical Oncology* **27**(10), 1621–1626. doi:10.1200/JCO.2008.16.6348
- Walling AM, Keating NL, Kahn KL, et al. (2016) Lower patient ratings of physician communication are associated with unmet need for symptom management in patients with lung and colorectal cancer. *Journal of Oncology Practice* **12**(6), e654–669. doi:10.1200/JOP.2015.005538
- Winters-Stone KM, Lyons KS, Bennett JA, et al. (2014) Patterns and predictors of symptom incongruence in older couples coping with prostate cancer. *Supportive Care in Cancer* **22**(5), 1341–1348. doi:10.1007/s00520-013-2092-0
- Yeager KA, Lee H, Bai J, et al. (2022) Congruence of pain perceptions between Black cancer patients and their family caregivers. *Supportive Care in Cancer* **30**(1), 543–553. doi:10.1007/s00520-021-06448-7