Analyzing Symptom Management Trials: The Value of Both Intention-to-Treat and Per-Protocol Approaches

Barbara Given, PhD, RN, FAAN, Charles W. Given, PhD, Alla Sikorskii, PhD, Mei You, MS, Ruth McCorkle, PhD, RN, FAAN, and Victoria Champion, DNS, RN, FAAN

Purpose/Objectives: Two analytical approaches are described for a randomized trial testing interventions for symptom management.

Design: To compare an intention-to-treat with a per-protocol approach.

Setting: Patients were accrued from six cancer centers.

Sample: 94 men and 140 women with solid tumors were accrued.

Methods: An intention-to-treat approach (as randomized) and per-protocol analyses (at least one symptom reaching threshold and one follow-up intervention) were compared. The analysis determines how each approach affects results. A two-arm, six-contact, eight-week trial was implemented. In one arm, nurses followed a cognitive behavioral protocol. In the second arm, a non-nurse coach referred patients to a symptom management guide.

Main Research Variables: Trial arm; summed severity scores; interference-based severity categories at intake, 10 weeks, and 16 weeks; site; and stage of cancer.

Findings: Each arm produced a reduction in severity at 10 and 16 weeks with no differences between arms. In the per-protocol analyses, symptoms reported at the first contact required more time to resolve. Older patients exposed to the nurse arm resolved in fewer contacts.

Conclusions: The intention-to-treat analyses indicated that both arms were successful but offered few insights into how symptoms or patients influenced severity. Per-protocol analyses (intervention and dose), when, and which strategies affected symptoms.

Implications for Nursing: Each analytical strategy serves a purpose. Intention-to-treat defines the success of a trial. Per-protocol analyses allow nurses to pose clinical questions about response and dose of the intervention. Nurses should participate in analyses of interventions to understand the conditions where interventions are successful.

Among patients with cancer undergoing chemotherapy, the occurrence and severity of symptoms are important indicators of adverse events as well as of compromises in the quality of patients' lives. National cooperative groups and community clinical oncology programs have focused on pharmacologic approaches to symptom management, whereas support for nonpharmacologic trials has been confined largely to the National Institutes of Health (NIH) R01 and R21 research project grant mechanisms (Buchanan, O'Mara, Kelaghan, & Minasian, 2005; Minasian et al., 2007; Sloan, Cella, & Hays, 2005). Cleeland (2007) defined cancer symptom burden as the sum of all symptoms reported by patients. He argued that reducing symptom burden is important, even if improved overall quality of life cannot be achieved.

The goals of this article are to present data from a two-arm trial to determine whether a nurse-directed cognitive behavioral approach to symptom management that tailored intervention strategies to patients around education, counseling, support, reframing, and rehearsal produced significant reductions in symptom severity, compared with an education information arm delivered by a non-nurse coach prepared with a master's degree in the social sciences.

In previous work, elaborate cognitive behavioral models proved significantly more effective in reducing symptom severity compared with conventional care alone (Given et al., 2004b; Miaskowski, Dodd, & Lee, 2004). However, when compared with alternative approaches, most notably education information strategies, cognitive behavioral models appeared no more effective (Jacobsen et al., 2002; Newell, Sanson-Fisher, & Savolainen, 2002; Yates et al., 2005). Therefore, a comparison of two approaches guided the design, implementation, and analysis of this trial.

To establish that a novel intervention reduces total symptom severity burden, a summary measure of symptom severity is required, and analysis must follow an intention-to-treat approach. However, a composite measure of symptom burden summarized as a single

outcome across multiple symptoms will fail to discern possible effects of selected characteristics on certain symptoms. Therefore, learning whether any symptoms respond more favorably or more quickly than others and under what conditions the responses occur is essential. In addition, for symptom-level assessments, measures of severity must be reliable over time, valid variations must be clinically meaningful, and differential symptom responses must not be a function of measurement error or assessment bias.

Based on Given et al. (2004b), where a cognitive behavioral arm was contrasted with conventional care, this article extends this line of investigation; compares the results of a two-arm, randomized trial of novel interventions for the management of 17 frequently occurring symptoms among patients with cancer who are undergoing chemotherapy; and specifies how each trial arm and patient characteristics relate to symptom response and time to response. The authors argue that intention-to-treat analyses of symptom management interventions are essential to identifying dimensions of trial efficacy. Based on the analysis of the sample of patients at randomization, regardless of their exposure, the intention-to-treat analyses seek to determine whether one arm outperforms the other at a specific trial endpoint, regardless of intervention dose and attrition. Therefore, intention-to-treat analysis suggests how, given similar patients and under similar conditions, the intervention strategies under testing might perform in the "real world" (Cleeland, 2007; Cohen, 1998; Hollis & Campbell, 1999; Montori & Guyatt, 2001; Piantadosi, 1997).

Intention-to-treat analyses summarize the impact of the interventions but offer few details regarding the pattern of responses or time to achieve response. However, per-protocol analyses specify the conditions under which interventions are more or less effective (Fairclough, 1997). For example, in this approach, patients who received only a portion of the interventions can be assessed, and analyses can describe the mechanisms of action through which patient characteristics, such as age, education, sites, or stages of cancer, might moderate the impact of specific intervention strategies on symptom responses (Baron & Kenny, 1986; Bellg et al., 2004; Czaja, Schulz, Lee, Belle, & REACH Investigators, 2003; Fairclough; Given et al., 2008; Kraemer, Frank, & Kupfer, 2006; Owen, Klapow, Hicken, & Tucker, 2001). In addition, given the different analytic approaches, the outcome measures differ as well. Intention-to-treat analyses require a single composite measure, such as the Symptom Distress Scale (McCorkle & Young, 1978), the Memorial Symptom Assessment Scale (Daut, Cleeland, & Flanery, 1983; Jeon, Given, Sikorskii, & Given, 2009; Portenoy et al., 1994), the Brief Pain and Brief Fatigue Inventories (Mendoza et al., 1999), or the M.D. Anderson Symptom Inventory (Cleeland et al., 2000), which are commonly used (Given et al., 2004b; Kirkova et al., 2006; Miaskowski et al., 2004; Trask, Paterson, Griffith, Riba, & Schwartz, 2003). Per-protocol approaches can reveal more with anchor-based symptom response outcomes that reveal how each symptom responds to the respective strategies delivered in each trial arm.

First, the authors present an intention-to-treat analysis, guided by the following question: When compared with a six-contact, eight-week intervention delivered by a non-nurse coach, does a six-contact, eight-week cognitive behavioral intervention delivered by oncology nurses produce a significantly greater reduction in symptom severity? This analysis allowed the authors to examine the effect of the respective arms on lowering overall symptom severity. Second, once the authors established the impact of each trial arm on reducing summed symptom severity, a per-protocol analysis allowed them to address the following questions: Using anchor-based measures of responses (based on cut points for mild, moderate, and severe levels unique to each symptom, does the six-contact, eight-week cognitive behavioral intervention delivered by nurses produce *more* symptom responses than the coach arm? Do certain symptoms respond more favorably to one intervention arm? Which arm produces these responses in the shortest time? In addition, the authors assessed the moderating effect of age on patients' responses to the interventions.

Materials and Methods

Identification of Study Participants

The sponsoring university's institutional review board (IRB), along with IRBs of two comprehensive cancer centers, one community cancer oncology program, and six hospital-affiliated community oncology centers, approved this research. Subcontracts with each center were completed, and nurses from the respective clinical trial offices were trained to implement the recruitment protocol. Patient eligibility criteria included being 21 years of age or older, having a diagnosis of a solid tumor cancer or non-Hodgkin lymphoma, undergoing a course of chemotherapy, being able to hear and to speak and read English, having a touchtone telephone, and having a family member whom the patient could call on for caregiving when assistance was needed. Participating patients and their family caregivers signed informed consent forms, and their sociodemographic information was entered into a Web-based tracking system. All patients entering the trial were screened twice weekly for six weeks to identify symptom severity using an automated voice response version of the M.D. Anderson Symptom Inventory (Kraemer et al., 2006). All patients scoring 3 or higher on severity of either pain or fatigue, or 2 or higher on both severity of pain and fatigue (range 0–10), entered the trial. Those not satisfying this screening criterion either entered

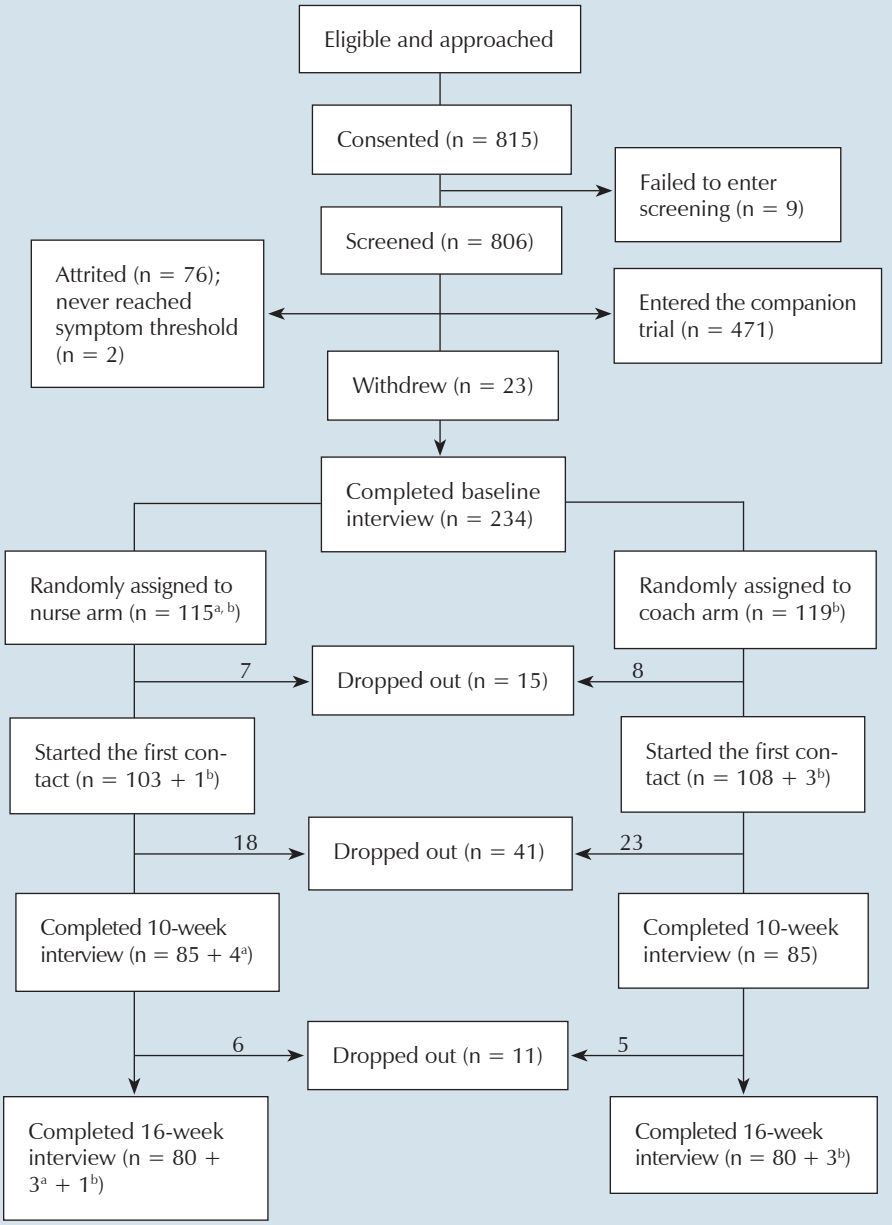
a companion trial or were sent a letter thanking them for participation but were not entered into the trial.

Upon entering the trial, eligible patients along with their family caregivers received an intake interview and a copy of the Symptom Management Guide (SMG) via mail and were randomized into either the Nurse Assisted Symptom Management (NASM) or the non-nurse Coach Assisted Symptom Management (CASM) arm by a computer minimization program that balanced the arms with respect to recruitment location and site of cancer (Taves, 1974). Interveners in both arms contacted patients via telephone. Patients in each trial arm received one intervention call each for the first four weeks, skipped week 5, were called week 6, skipped week 7, and received a final intervention call on week 8. Family members received calls at weeks 1, 4, and 8. At 10 and 16 weeks, outcome data were obtained on both members of the dyad through a second and third interview. The analyses presented here focus on the patients. Figure 1 summarizes the number of enrolled and attritional patients at each trial point, the number of cases meeting entry criteria, and the number analyzed.

Power calculations for this trial were based on an earlier trial where, when compared with conventional care alone, a 10-contact, 20-week NASM produced a 15-point reduction in summed severity following five contacts at 10 weeks (Given et al., 2004b). Based on estimates from a pilot study, the authors expected that CASM would reduce symptom severity by five points. Therefore, the trial was powered to detect a difference of 10 points following six intervention contacts delivered over eight weeks with outcome observations at 10 weeks. Such differences between arms corresponded to the effect size of 0.36 and, to be detected as statistically significant, required 122 patients per arm (175 per arm to offset attrition). After nearly 66% of the patients had been exposed to the intervention, the observed effect size for the 10-week endpoint was assessed. Virtually no differences in summed symptom severity scores were found between arms (effect sizes of 0.03) at 10 weeks. Because effect sizes of this magnitude are not clinically important, patient accrual was suspended (Cohen, 1998; Sloan et al., 2007). Comparisons between baseline and 10- and 16-week endpoints for each arm revealed large effect sizes (0.82 and 0.59 for the NASM, 0.65 and 0.72 for the CASM). Therefore, although virtually no differences were observed between the arms, each arm had improvements over baseline of a large magnitude.

Trial Arms

Extending past work (Given et al., 2004a; Rawl et al., 2002), this trial compared a nurse-directed cognitive behavioral approach with an education information intervention. The conceptual work underpinning this intervention has been previously reported (Given et al.,



[a] Patients skipped intervention (four in nurse arm); three completed 16-week interview.
[b] Patients skipped 10-week interview (one in nurse arm, three in coach arm).

Figure 1. Flowchart of Accrual, Enrollment, and Attrition

2004a, 2004b; Rawl et al.). Seventeen symptoms were targeted: fatigue, pain, dyspnea, insomnia, distress, nausea, fever, difficulty remembering, lack of appetite, dry mouth, vomiting, numbness and tingling, diarrhea, cough, constipation, weakness, and alopecia. Following National Comprehensive Cancer Network ([NCCN], 2006) guidelines, patients who rated severities of symptoms at a score of four or higher (threshold) at each contact received strategies to manage those symptoms. Symptoms at a score of seven or higher were deemed urgent according to guidelines, and patients were encouraged to call the oncology office if the symptoms did not improve.

For patients assigned to the NASM arm, nurses delivered up to four strategies for each symptom supplemented with references to the SMG. At each subsequent contact, the nurse inquired whether the recommended strategy was tried and whether it was helpful in managing the symptom. When strategies were not tried, or tried but found not to be helpful, patients were counseled as to how a strategy might fit into their daily activities or they were offered different strategies. Successful intervention strategies were retained and the nurse reinforced the importance of continuing to use the strategies.

In the CASM arm, a trained coach who followed a script assessed patients regarding the severity of each of the 17 symptoms. Coaches were persons with education in the social sciences and their training included simulations, rehearsals, and practice sessions to ensure that their role was simply to encourage use of the SMG and not to give advice. Patients' responses were recorded on the computer protocol. For those symptoms rated at a score of four or higher, the coach directed patients to read sections of the SMG. For each symptom, a section informed patients about management strategies. At each subsequent call, the coach first asked patients whether or not they had read the SMG, tried the recommended strategies, and, if so, found them to be helpful in lowering the severity of that symptom. When all symptoms above threshold at the previous contact were evaluated, the coach then reviewed the current severity of all symptoms. Therefore, the trial arms were identical in terms of the symptoms addressed, how they were scored, the number of contacts, and use of the SMG. They differed only in how strategies for symptom management were presented to the patient and how they were modified with regard to patient responses to the information.

Five nurses were employed to implement the cognitive behavioral arm. To train and to ensure fidelity of this arm, the authors confirmed that all nurses were certified by the Oncology Nursing Certification Corporation (an affiliate of the Oncology Nursing Society [ONS]), possessed at least two years' experience in oncology practices (none of which were study sites), and had received 20 hours of training using a manual derived from the Web-based protocol they would be required to follow during the intervention. The authors used simulated patients for rehearsal and practice sessions prior to implementing the protocol. Nurses completed continuing education sessions on symptom management from the ONS Web site. This protocol focused on assessing the severity and interference of the 17 symptoms. All symptoms scored at four or higher on severity were automatically transferred to the "plan of care" in the Web-based data file. Nurses were trained to work with patients to select up to four symptoms to address at each contact. Following symptom selection, nurses then provided patients with evidence-based strategies specific to that symptom. The authors divided these strategies into four areas: education and information, support, communication, and reframing. Nurses could assign up to four strategies for each symptom. At each subsequent contact, the authors assessed patients to determine whether they tried each strategy for the respective symptom and, if so, the extent to which it helped to relieve the symptom. Untested or ineffective strategies were altered and replaced with new ones through negotiations with patients. All symptoms were assessed again at each contact and, for those above threshhold, new strategies were agreed on and assigned to patients. A senior project nurse (a nurse practitioner with cancer experience) reviewed all nurses' documentation in the Web-based plan of care. The senior project nurse analyzed tapes of the nurse-patient interactions at each intervention and addressed deviations with nurses in biweekly meetings or immediately if necessary.

For the education information arm, the authors trained two coaches to deliver the intervention. The coaches assessed the 17 symptoms at each contact, and for the symptoms above threshold, assigned the patient to read sections in the SMG describing management strategies for that symptom. At subsequent contacts, the coach first evaluated patients' use of the SMG for each symptom above threshold at the prior contact: Were the strategies tried and, if so, were they helpful in managing the symptom? To ensure fidelity of the system, a senior project nurse monitored this version of the Web-based strategy and documentation of the interaction, in addition to monitoring thorough reviews of the recorded interventions between the patient and the coach.

Measures

The authors obtained sociodemographic characteristics from the patients' medical records, entered them into the tracking system, and confirmed during the baseline interview. The authors recorded age in years and categorized as 54 years or younger, 55–64 years, and 65 years or older. The authors assessed comorbid conditions at baseline using a 13-item questionnaire: Patients

Table 1. Characteristics of Patients by Trial Arm

Characteristic	Nurse Arm (N = 115) n	Nurse Arm (N = 115) %	Coach Arm (N = 119) n	Coach Arm (N = 119) %	Chi-Square p
Cancer site					0.95
Breast	27	24	28	24	
Lung	37	32	40	34	
Other	51	44	51	42	
Recurrence					0.42
Yes	50	43	44	37	
No	63	55	69	58	
Missing	2	2	6	5	
Metastatic					0.07
Yes	85	74	74	62	
No	30	26	44	37	
Missing	–	–	1	1	
Gender					0.83
Male	47	41	47	40	
Female	68	59	72	60	
Cancer stage					0.63
Early	15	13	18	15	
Late	100	87	100	84	
Missing	–	–	1	1	
Age (years)					1
54 or younger	41	36	43	36	
55–64	39	34	40	34	
65 or older	35	30	36	30	
Comorbidity					0.39
0–1	43	37	51	43	
2 or more	72	63	68	57	
CESD scale					0.12
Less than 16	79	69	70	59	
More than 16	36	31	49	41	

CESD—Center for Epidemiologic Studies Depression

were asked whether a doctor had ever told them they had such conditions as diabetes, high blood pressure, and other chronic diseases (Katz, Chang, Sangha, Fossel, & Bates, 1996).

Patients scored severity of 17 symptoms on a scale ranging from absence (0) to the worst severity possible (10) at baseline, 10- and 16-week interviews, and at each intervention contact (telephone call). Scores reported at baseline and 10- and 16-week interviews were summed across symptoms creating an index of severity ranging from 0–170, and these scores were used in the intention-to-treat analysis. This measure was based on NCCN guidelines (2006) used to assess symptom severity and used in previous randomized, controlled trials. For the per-protocol analysis, the authors drew on the status of each symptom from each of the six intervention contacts. Based on prior work using data obtained on symptoms from this and a companion trial, cut points for each symptom were identified based on comparing differences in interference scores associated with successive increases in severity (Baron & Kenny, 1986; Portenoy et al., 1994). For example, for pain and fatigue, the mild category corresponds to a severity score of 1, the moderate category corresponds to scores of 2–4, and scores of 5–10 fall into the severe category. For insomnia and peripheral neuropathy, the mild category is 1–3, the moderate is 4–6, and severe is 7–10. The cut points were then tested and found to differentiate consistently with the levels of interference associated with mild, moderate, and severe scores at successive intervention contacts. In establishing the cut points, as well as their stability, the authors observed no differences between this and the Given et al. (2008) companion trial.

Based on these sustained differences in interference at the six contacts, anchor-based definitions of response included severity categories at symptom onset (date of the contact when the symptom first reached moderate or severe) and at the last contact completed by a patient. The authors classified transitions from severe to moderate or mild (or none when the symptom was no longer present) or from moderate to mild (or none) between onset and last contact completed as responses for each symptom. Nonresponses are defined as transitions from severe to severe or from moderate to moderate or severe (where the onset and endpoint observations fail to define a downward shift). Nonresponders' time to response was treated as censored; for responders, time in days from the symptom onset to the date of the first contact when patients reported sustained improvement (e.g., going from moderate to mild and staying mild for the remaining contacts) was defined as "time to response."

For the per-protocol analysis, the total number of symptoms that reached moderate or severe levels during the six contacts was determined and dichotomized at the median as 6 or below versus greater than 6. To assess time of onset and time to response, onset time for each symptom was classified into two categories: first contact (symptom present prior to any intervention) and second or later contact.

Intention-to-Treat Analysis

All patients were analyzed as randomized, regardless of their adherence to or attrition from the intention protocol. Baseline equivalence of trial arms was established with chi-square and t tests. Attrition analysis was

Table 2. Symptom Severity of Patients by Trial Arm

Study Measure Point	Nurse Arm N	Nurse Arm $\bar{X}$	Nurse Arm SD	Coach Arm N	Coach Arm $\bar{X}$	Coach Arm SD	T-Test p
Baseline	115	40.64	21.49	119	39.64	23.34	0.73
10 weeks	89	21.82	17.73	85	21.05	16.62	0.77
16 weeks	84	20.46	20.46	83	19.29	16.47	0.12

conducted to ensure absence of bias in the analysis of outcomes. Baseline symptom severity was compared by trial arm for patients who were lost and remained in the trial. Demographic variables were compared by arm for patients who withdrew from the trial between baseline and10-week interview and between baseline and 16-week interview.

Linear mixed effects model implemented in SAS® (v.9.1) related symptom severity at 10 and 16 weeks to baseline symptom severity, trial arm, time (10 or 16 weeks), patient comorbid conditions, metastatic versus local disease, age, and age by trial arm interaction. The adjusted means for age categories were compared by trial arm. Residuals were examined to evaluate model fit and outliers. To assess the changes from baseline to 10 weeks and from baseline to 16 weeks, matched paired t tests were performed for each arm.

Per-Protocol Analysis

Only patients who had at least one symptom with severity reaching moderate or severe and with a follow-up contact were included in this analysis. Patients not analyzed were those who never reported a symptom above mild because no possibility existed for improvement and those who reached moderate or severe levels for the first time at the last contact because no opportunity existed to assess the impact of interventions delivered. To analyze response versus nonresponse across multiple symptoms within a patient, the authors used Generalized Estimating Equations (GEE) model with compound symmetry correlation structure. Covariates included trial arms, comorbid condition, onset time of each symptom, metastatic status, age, and its interaction with trial arms. Odds ratios and their 95% confidence intervals were estimated for trial arms, interaction with age, and other covariates of interest. The GEE model was implemented the GENMOD procedure in SAS.

Assessment of time-to-symptom response (in days between contacts) was carried out with a marginal Cox proportional hazard model implemented in TPHREG procedure in SAS with the same covariates as employed in the response analysis. The Lee, Wei, and Amato (1992) method was used for aggregating multiple symptoms at the patient level. A robust sandwich covariance accounted for the intracluster dependence (symptoms clustered within a patient). Hazard ratios and the 95% confidence intervals were evaluated for categorical variables and the trial arm effect within different age groups.

Results

Table 1 summarizes the characteristics of patients by trial arm. No differences were found between arms among patients lost from the trial according to gender, cancer site or stage, comorbid conditions, age, depression, or metastatic versus local disease. Second, the logistic analysis indicated that patients with greater symptom severity were significantly more likely to be lost by the 16-week endpoint. Baseline summed severity for lost patients was 44.8 and 52.3, respectively, for the nurse and coach arm; this was contrasted with baseline scores for those who remained in the trial of 38.7 and 33.8. However, no differences existed between arms according to the severity scores of patients who remained or were lost. Therefore, the internal validity of the trial remained, but the external validity (generalizability) was threatened.

For the intention-to-treat analysis, the t tests for group comparisons at baseline, 10 weeks, and 16 weeks are summarized in Table 2. At 10 weeks, the arms produced between a 14- and 17-point reduction in summed symptom severity over baseline with virtually no change at 16 weeks, suggesting a sustained effect following the 10-week trial endpoint. However, at neither endpoint were differences observed between the two arms. Table 3 contains the means for the summed symptom severity at 10 and 16 weeks adjusted for metastatic status and comorbidity overall by arm and by arm within the age categories. Again, the authors observed no significant differences; therefore, the intention-to-treat analysis

Table 3. Means of Summed Symptom Severity at 10 and 16 Weeks Adjusted for Metastatic Status and Comorbidity

Covariate	Least Square $\bar{X}$	SE	p
Intervention trial arm			
Nurse arm at 10 weeks	19.98	1.77	0.67
Coach arm at 10 weeks	21.01	1.74	
Nurse arm at 16 weeks	22.03	1.98	0.31
Coach arm at 16 weeks	19.28	1.94	
Interaction of age in years by trial arm			
54 or younger (nurse arm at 10 weeks)	22.1	2.81	0.33
54 or younger (coach arm at 10 weeks)	18.14	2.9	
55–64 (nurse arm at 10 weeks)	16.92	2.89	0.24
55–64 (coach arm at 10 weeks)	21.83	3.01	
65 and older (nurse arm at 10 weeks)	20.93	3.3	0.62
65 and older (coach arm at 10 weeks)	23.07	3.14	
54 or younger (nurse arm at 16 weeks)	23.18	3.2	0.18
54 or younger (coach arm at 16 weeks)	17.06	3.23	
55–64 (nurse arm at 16 weeks)	19.45	3.24	0.96
55–64 (coach arm at 16 weeks)	19.69	3.45	
65 and older (nurse arm at 16 weeks)	23.47	3.66	0.62
65 and older (coach arm at 16 weeks)	21.07	3.42	

SE—standard error

indicates that the more elaborate cognitive behavioral intervention delivered by nurses was no more successful in lowering symptom severity than was the intervention delivered by coaches. Although such changes could have occurred by chance, differences of this magnitude between intake and outcome would occur less than 5 times out of 100. In addition, some patients had concluded their courses of chemotherapy. However, comparisons in each group between those continuing and those who had concluded their chemotherapy indicate that no differences in symptom severity existed.

The per-protocol analyses examined the differences between trial arm in terms of symptom response during six intervention contacts. Table 4 describes each symptom, the moderate and severe categories using the interference-based severity cut points described, and the number of patients with each symptom who responded or did not respond according to trial arm.

No differences in symptom responses by trial arm were noted (see Tables 5 and 6). Symptoms with onset at the second or later intervention contact were more likely to resolve than those identified as moderate or severe at first contact. Among older patients, the nurses appeared to be more successful than the coaches, but trial arm by age interaction did not reach statistical significance.

Finally, when the time in days needed to produce a response was assessed, symptoms that rose above threshold at the second or later contacts had a shorter time to response and, when compared with the coach arm, nurses were able to produce significantly shorter time to responses among patients who were 65 years of age or older.

Discussion

The intention-to-treat and per-protocol analyses indicate that an education and information intervention delivered by a coach who assesses symptom severity and refers patients with symptoms above threshold to an SMG produces reductions in symptom severity at 10 weeks and sustains those reductions at 16 weeks, comparable to a more complex cognitive behavioral intervention delivered by trained cancer nurses. These findings are consistent with other trials and trial summaries (Jacobsen et al., 2002; Newell et al., 2002). In addition, the coach arm was able to produce these outcomes in just over 18 minutes per contact, whereas the nurses required 42 minutes on average across each of the contacts to achieve similar results.

The intention-to-treat analysis did not attempt to impute missing cases, which, in a full intention-to-treat

Table 4. Ranges of Moderate and Severe Categories of Symptoms and Distributions of Patient Responses by Trial Arm

Severity of Symptom Cut Points	Nurse Arm				Coach Arm			
	No Response		Response		No Response		Response	
	n	%	n	%	n	%	n	%
Anxiety								
Moderate (4–5)	4	13	26	87	8	26	23	74
Severe (6–10)	2	12	14	88	4	21	15	79
Appetite								
Moderate (4–5)	6	27	16	73	6	21	23	79
Severe (6–10)	5	22	18	78	7	28	18	72
Constipation								
Moderate (4–6)	5	33	10	67	5	23	17	77
Severe (7–10)	1	9	10	91	3	25	9	75
Cough								
Moderate (3–4)	7	37	12	63	7	39	11	61
Severe (5–10)	2	20	8	80	3	50	3	50
Depression								
Moderate (2–3)	3	14	19	86	7	27	19	73
Severe (4–10)	2	8	22	92	6	21	22	79
Diarrhea								
Moderate (4–5)	2	25	6	75	2	9	20	91
Severe (6–10)	1	8	12	92	3	50	3	50
Dry mouth								
Moderate (5–8)	8	32	17	68	7	25	21	75
Severe (9–10)	–	–	3	100	–	–	3	100
Dyspnea								
Moderate (3–6)	14	44	18	56	18	46	21	54
Severe (7–10)	1	25	3	75	1	20	4	80
Fatigue								
Moderate (2–4)	22	50	22	50	27	64	15	36
Severe (5–10)	16	31	36	69	21	38	34	62
Vomiting								
Moderate (4–6)	5	25	15	75	1	5	19	95
Severe (7–10)	–	–	8	100	2	29	5	71
Pain								
Moderate (2–4)	11	29	27	71	16	38	26	62
Severe (5–10)	9	38	15	62	7	29	17	71
Peripheral neuropathy								
Moderate (4–7)	11	42	15	58	4	17	20	83
Severe (8–10)	2	50	2	50	2	29	5	71
Remembering								
Moderate (2–4)	6	17	30	83	7	22	25	78
Severe (5–10)	3	21	11	79	3	38	5	62
Sleep disturbance								
Moderate (4–6)	6	21	22	79	11	25	33	75
Severe (7–10)	2	11	16	89	4	18	18	82
Weakness								
Moderate (3–4)	8	25	24	75	15	41	22	59
Severe (5–10)	9	31	20	69	16	44	20	56

Table 5. Adjusted Odds Ratios of Symptom Responses Aggregated at Patient Level

Covariate	Level	Reference Level	Adjusted Odds Ratio (95% CI)	p
Trial arm	Nurse	Coach	–	0.11
Comorbidity	2 or more	0~1	0.77 (0.55, 1.09)	0.14
Onset	Second contact or later	First contact	2.4 (1.78, 3.23)	<0.01
Metastatic	Yes	No	0.67 (0.47, 0.94)	0.02
Age (years)	55–64 65 or older	54 or younger	– –	<0.01 0.03
Interaction of age by trial arm	54 or younger (nurse) 55–64 (nurse) 65 or older (nurse)	54 or younger (coach) 55–64 (coach) 65 or older (coach)	0.71 (0.42, 1.2) 1.57 (0.91, 2.73) 1.72 (0.94, 3.15)	0.2 0.11 0.08

CI–confidence interval

design, might have altered findings. Those with at least one postintervention interview (either at 10 or 16 weeks or both) were included in the analysis via linear mixed effects model. Imputation of outcomes for those who only completed baseline interview was not performed because of potential issues; techniques for imputations on covariates are well developed and are deemed reliable, whereas methods for imputation of outcomes may not perform well (Crawford, Tennstedt, & McKinlay, 1995). The attrition analysis revealed no differences by arm according to patient characteristics or severity scores at 10 or 16 weeks. This suggests that, although the trial may have lost generalizability, internal validity remained strong.

Intention-to-treat analyses meet scientific requirements for determining whether novel interventions produce significantly different results. They do not, however, elaborate on the complexity of findings, particularly as they pertain to the testing of symptom management interventions. Summed severity scores mask possible time to response or symptom-specific responses. In the per-protocol analysis, the authors aggregated symptoms within patients using appropriate analytical techniques and found that symptoms identified by patients at the initial contact are most troubling to them. A higher number of intervention strategies or more focused intervention strategies are needed to achieve a response, as well as more time to produce a response. Finally, regardless of the intervention arm, older patients appear to respond more favorably to the nurse, and when the analysis is focused on the time to response, the nurse arm produces those responses in a significantly shorter period of time.

From the perspective of clinicians, the intention-to-treat analysis indicates that the more complex and costly nurse-delivered approach appears no better than an inexperienced intervener who assesses patients' symptoms, refers them to the SMG, and evaluates patients' use of the material. The per-protocol approach successfully specified the conditions under which the nurse intervention produces symptom response in a shorter period of time. Each analytical model offers important information. Intention to treat defined the outcome, the extent to which it was generalizable, and the per-protocol analysis elaborated at the symptom level important trial arm differences, which enable better specification and clinical relevance of trial effectiveness. The interference-based severity cut points that differentiate severe from moderate and mild and moderate from mild are important advances in the measurement of symptoms as they permit researchers to conduct analyses based on responses (severe to moderate or mild) defined for each symptom. Based on this approach, the authors argue that all symptom management trials should adopt both an intention-to-treat as well as a per-protocol approach so that efficacy and effectiveness can be presented to researchers and practitioners.

Several limitations in this trial must be acknowledged. First, the intention-to-treat analysis did not attempt to

Table 6. Adjusted Hazard Ratios of Time to Response (in Days) in Symptoms Aggregated at the Patient Level

Covariate	Level	Reference Level	Adjusted Odds Ratio (95% CI)	p
Trial arm	Nurse	Coach	–	0.33
Comorbidity	2 or more	0~1	0.85 (0.71, 1.02)	0.08
Onset	Second contact or later	First contact	2.01 (1.73, 2.33)	<0.01
Metastatic	Yes	No	0.88 (0.73, 1.06)	0.16
Age (years)	55–64 65 or older	54 or younger	– –	0.04 0.01
Interaction of age by trial arm	54 or younger (nurse) 55–64 (nurse) 65 or older (nurse)	54 or younger (coach) 55–64 (coach) 65 or older (coach)	0.87 (0.71, 1.07) 1.11 (0.87, 1.4) 1.52 (1.2, 1.93)	0.19 0.4 <0.01

CI—confidence interval

impute missing data related to patient attrition after baseline interview. The analyses demonstrate the impact of attrition on generalizability, and the authors agree that it did not influence the integrity of the trial analysis. Because of the lack of differences between arms, the trial was discontinued because it would have required more than 1,000 cases to detect differences that were not clinically or practically important, despite the fact that the nurse arm spent on average 44 minutes per patient whereas the coach arm averaged 19 minutes of direct patient contact.

In conclusion, this research demonstrated the benefit of presenting both an intention-to-treat and a per-protocol analysis. Future trials should present both approaches and establish not only that the interventions reduce symptom burden, but also where and for whom those burdens are lightened and to explore how these differences occur at the level of the symptoms or the time required to produce responses.

Barbara Given, PhD, RN, FAAN, is a university distinguished professor in the College of Nursing, Charles W. Given, PhD, is a professor in the Department of Family Medicine, Alla Sikorskii, PhD, is an assistant professor in the College of Nursing and Department of Statistics, and Mei You, MS, is a statistical assistant in the College of Nursing, all at Michigan State University in East Lansing; Ruth McCorkle, PhD, RN, FAAN, is the Florence Wald Professor of Nursing in the School of Nursing at Yale University in New Haven, CT; and Victoria Champion, DNS, RN, FAAN, is an associate dean for research in the School of Nursing and a director of cancer control in the Simon Cancer Center at Indiana University in Indianapolis. This research was funded by a grant from the National Cancer Institute (#CA 79280). Charles W. Given can be reached at givenc@msu.edu, with copy to editor at ONFEditor@ons.org. (Submitted August 2008. Accepted for publication March 7, 2009.)

Digital Object Identifier: 10.1188/09.ONF.E293-E302

References

Baron, R.M., & Kenny, D.A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic and statistical considerations. *Journal of Personality and Social Psychology, 51*(6), 1173–1182.

Bellg, A.J., Borrelli, B., Resnick, B., Hecht, J., Minicucci, D.S., Ory, M., et al. (2004). Enhancing treatment fidelity in health behavior change studies: Best practices and recommendations from the NIH Behavior Change Consortium. *Health Psychology, 23*(5), 443–451.

Buchanan, D.R., O'Mara, A.M., Kelaghan, J.W., & Minasian, L.M. (2005). Quality-of-life assessment in the symptom management trials of the National Cancer Institute–Supported Community Clinical Oncology Program. *Journal of Clinical Oncology, 23*(3), 591–598.

Cleeland, C.S. (2007). Symptom burden: Multiple symptoms and their impact as patient-reported outcomes. *Journal of the National Cancer Institute. Monographs, 37,* 16–21.

Cleeland, C.S., Mendoza, T.R., Wang, X.S., Chou, C., Harle, M.T., Morrissey, M., et al. (2000). Assessing symptom distress in cancer patients: The M.D. Anderson Symptom Inventory. *Cancer, 89*(7), 1634–1646.

Cohen, J. (1998). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NY: Lawrence Erlbaum Associates.

Crawford, S.L., Tennstedt, S.L., & McKinlay, J.B. (1995). A comparison of analytic methods for non-random missingness of outcome data. *Journal of Clinical Epidemiology, 48*(2), 209–219.

Czaja, S.J., Schulz, R., Lee, C.C., Belle, S.H., & REACH Investigators. (2003). A methodology for describing and decomposing complex psychosocial and behavioral interventions. *Psychology and Aging, 18*(3), 385–395.

Daut, R.L., Cleeland, C.S., & Flanery, R.C. (1983). Development of the Wisconsin Brief Pain Questionnaire to assess pain in cancer and other diseases. *Pain, 17*(2), 197–210.

Fairclough, D.L. (1997). Summary measures and statistics for comparison of quality of life in a clinical trial of cancer therapy. *Statistics in Medicine, 16*(11), 1197–1209.

Given, B., Given, C.W., Sikorskii, A., Jeon, S., McCorkle, R., Champion, V., et al. (2008). Establishing mild, moderate, and severe scores for cancer-related symptoms: How consistent and clinically meaningful are interference-based severity cut-points? *Journal of Pain and Symptom Management, 35*(2), 126–135.

Given, C., Given, B., Rahbar, M., Jeon, S., McCorkle, R., Cimprich, B., et al. (2004a). Does a symptom managment intervention affect depression among cancer patients: Results from a clinical trial. *Psycho-Oncology, 13*(11), 818–830.

Given, C., Given, B., Rahbar, M., Jeon, S., McCorkle, R., Cimprich, B., et al. (2004b). Effect of a cognitive behavioral intervention on reducing symptom severity during chemotherapy. *Journal of Clinical Oncology, 22*(3), 507–516.

Hollis, S., & Campbell, F. (1999). What is meant by intention to treat analysis? Survey of published randomized controlled trials. *BMJ, 319*(7211), 670–674.

Jacobsen, P.B., Meade, C.D., Stein, K.D., Chirikos, T.N., Small, B.J., & Ruckdeschel, J.C. (2002). Efficacy and costs of two forms of stress management training for cancer patients undergoing chemotherapy. *Journal of Clinical Oncology, 20*(12), 2851–2862.

Jeon, S., Given, C.W., Sikorskii, A., & Given, B. (2009). Do interference-based cut-points differentiate mild, moderate, and severe levels of 16 cancer-related symptoms over time? *Journal of Pain and Symptom Management, 37*(2), 220–332.

Katz, J.N., Chang, L.C., Sangha, O., Fossel, A.H., & Bates, D.W. (1996). Can comorbidity be measured by questionnaire rather than medical record review? *Medical Care, 34*(1), 73–84.

Kirkova, J., Davis, M.P., Walsh, D., Tiernan, E., O'Leary, N., LeGrand, S.B., et al. (2006). Cancer symptom assessment instruments: A systematic review. *Journal of Clinical Oncology, 24*(9), 1459–1473.

Kraemer, H.C., Frank, E., & Kupfer, D.J. (2006). Moderators of treatment outcomes: Clinical, research and policy importance. *JAMA, 296*(10), 1286–1289.

Lee, E. Wei, L.J., & Amato, D. (1992). Cox-type regression analysis for large numbers of small groups of correlated failure time observations. In J.P. Klein and P.K. Goel (Eds.), *Survival analysis: State of the art, NATO ASI Series* (Vol. 211). Amsterdam, Netherlands: Kluwer Academic Publishers.

McCorkle, R., & Young, K. (1978). Development of a symptom distress scale. *Cancer Nursing, 1*(5), 373–378.

Mendoza, T.R., Wang, X.S., Cleeland, C.S., Morrissey, M., Johnson, B.A., Wendt, J.K., et al. (1999). The rapid assessment of fatigue severity in cancer patients: Use of the Brief Fatigue Inventory. *Cancer, 85*(5), 1186–1196.

Miaskowski, C., Dodd, M., & Lee, K. (2004). Symptom clusters: The new frontier in symptom management research. *Journal of the National Cancer Institute. Monographs, 32,* 17–21.

Minasian, L.M., O'Mara, A.M., Reeve, B.B., Denicoff, A.M., Kelaghan, J., Rowland, J.H., et al. (2007). Health-related quality of life and symptom management research sponsored by the National Cancer Institute. *Journal of Clinical Oncology, 25*(32), 5128–5132.

Montori, V.M., & Guyatt, G.H. (2001). Intention-to-treat principle. *Canadian Medical Association Journal, 165*(10), 1339–1341.

National Comprehensive Cancer Network. (2006). Clinical Practice Guidelines in Oncology™. Retrieved July 15, 2008, from http://www.nccn.org/professionals/physician_gls/f_guidelines.asp

Newell, S.A., Sanson-Fisher, R.W., & Savolainen, N.J. (2002). Systematic review of psychological therapies for cancer patients: Overview and recommendations. *Journal of the National Cancer Institute, 94*(8), 558–584.

Owen, J.E., Klapow, J.C., Hicken, B., & Tucker, D. (2001). Psychosocial interventions for cancer: Review and analysis using a three-tiered outcomes model. *Psycho-Oncology, 10*(3), 218–230.

Piantadosi, S. (1997). Overviews/meta-analyses. In S. Piantadosi (Ed.), *Clinical trials: A methodologic perspective* (p. 261). New York: Wiley-Interscience.

Portenoy, R.K., Thaler, H.T., Kornblith, A.B., Lepore, J.M., Friedlander-Klar, H., Kiyasu, E., et al. (1994). The Memorial Symptom Assessment Scale: An instrument for the evaluation of symptom prevalence, characteristics and distress. *European Journal of Cancer, 30A*(9), 1326–1336.

Rawl, S.M., Given., B.A., Given, C.W., Champion, V.L., Kozachik, S.L., Kozachik, S.L. et al. (2002). Intervention to improve psychological functioning for newly diagnosed patients with cancer. *Oncology Nursing Forum, 29*(6), 967–975.

Sloan, J.A., Berk, L., Roscoe, J., Fisch, M.J., Shaw, E.G., Wyatt, G., et al. (2007). Integrating patient-reported outcomes into cancer symptom management clinical trials supported by the National Cancer Institute–Sponsored Clinical Trials Networks. *Journal of Clinical Oncology, 25*(32), 5070–5077.

Sloan, J.A., Cella, D., & Hays, R.D. (2005). Clinical significance of patient-reported questionnaire data: Another step toward consensus. *Journal of Clinical Epidemiology, 58*(12), 1217–1219.

Taves, D.R. (1974). Minimization: A new method of assigning patients to treatment and control groups. *Clinical Pharmacology and Therapeutics, 15*(5), 443–453.

Trask, P.C., Paterson, A.G., Griffith, K.A., Riba, M.B., & Schwartz, J.L. (2003). Cognitive-behavioral intervention for distress in patients with melanoma. *Cancer, 98*(4), 854–864.

Yates, P., Aranda, S., Hargraves, M., Mirolo, B., Clavarino, A., McLachlan, S., et al. (2005). Randomized controlled trial of an educational intervention for managing fatigue in women receiving adjuvant chemotherapy for early-stage breast cancer. *Journal of Clinical Oncology, 23*(25), 6027–6036.