Cancer and its treatment produce multiple symptoms that significantly distress patients and impair function. Symptoms caused by treatment may delay treatment or lead to premature treatment termination, and residual treatment-related symptoms often complicate posttreatment rehabilitation. When treatment is no longer possible, symptom control becomes the focus of cancer care. Patient ratings of symptom severity and impact are important patient-reported outcomes (PROs) in cancer clinical trials and comprise a subset of a larger domain of PROs generally referred to as health-related quality of life (HRQOL). Symptoms rarely occur in isolation; rather, there is now ample evidence that symptoms frequently occur in clusters. The impact of these multiple symptoms upon the patient can be described as “symptom burden,” a concept that encompasses both the severity of the symptoms and the patient’s perception of the impact of the symptoms. The distress caused by symptoms is a subject of much investigation, and several validated measures of the severity and impact of multiple symptoms are now available. Symptom measures are generally brief, thereby reducing respondent burden, and can be administered repeatedly during a trial to give a relatively fine-grained picture of the patient’s status across time. In many instances, information on trial-related changes in symptom burden, or comparison of symptom burden between arms in a clinical trial, may provide sufficient self-report data for clinical trial consumers (patients, clinicians, and regulators) to make treatment choices or to evaluate new therapies, without measuring other HRQOL domains.

J Natl Cancer Inst Monogr 2007;37:16–21

On February 2, 2006, the U.S. Food and Drug Administration issued draft guidance for industry on Patient-Reported Outcome Measures: Use in Medical Product Development to Support Labeling Claims (1). This publication has led to wide discussion about domains of patient-reported outcomes (PROs) that should be considered as endpoints in clinical trials and for labeling purposes. Reducing the severity and impact of patient-reported symptoms is naturally an endpoint for symptom-focused interventions in clinical trials, and comparing treatment-related symptoms would provide an additional benchmark for appraising various cancer treatments. Because clinicians and patients commonly face choices among treatments that are similarly effective for tumor control and prolonging survival, differences in the patient’s status during the survival period have become critical variables in making final, individualized treatment choices and in developing new therapies. Including PROs as important measures of differences among treatments is paramount for effectively evaluating toxicity and quality of survival.

PROs can take a variety of forms, including measures of differences in symptom severity and impact and measures of health-related quality of life (HRQOL). A consensus among health researchers is that HRQOL is a multidimensional construct composed of at least four dimensions: physical function (i.e., daily activities, self-care), psychologic function (i.e., emotional or mental state, mood), social role function (i.e., social interactions, family dynamics), and disease or treatment symptoms (i.e., pain, nausea) (2). Symptom reports represent a subset of HRQOL. In most conceptualizations of HRQOL, symptoms are viewed as the patient report most proximal to the disease process and are thus potentially causative of variation in the more abstract HRQOL components such as well-being, perception of daily functioning, global impressions of the impact of treatment on daily life, satisfaction with treatment, and perception of overall health status (3).

Most HRQOL measures therefore include components that evaluate the severity of some symptoms. Commonly used HRQOL measures, including the Medical Outcomes Study Short Form-36 (4), the Functional Assessment of Cancer (5), and the European Organisation for Research and Treatment Of Cancer Quality of Life Questionnaire (EORTC QLQ-C30) (6), address major symptoms such as pain, depression, fatigue, and nausea. In the EORTC QLQ-C30, most of the items (18 of 30) are self-reported symptoms. Other HRQOL scales ask about social and role function and concerns about social support, domains that are more distal to treatment effects.

As patient outcome variables become more important relative to “hard” endpoints such as survival or time to recurrence, choices about which PRO measures are most meaningful and how to interpret them to patients, health care professionals, and policy makers need to be made. The usefulness of most of these PRO measures will depend ultimately on how patients report their perceived present status relative to how they recall their status before they got their disease and were treated for it. This review will address the status of symptom measures as outcomes for clinical research and compare multiple-symptom outcomes with those that encompass a larger set of HRQOL domains.
Symptoms as a Focus

Therapies that have symptom control as a focus have become more prominent in the past few years (7), with several factors playing a role in this development. Increasing demands from patients that they be more comfortable and functional are evident in the media and in the clinic. Symptom PROs have been accepted more frequently by regulatory agencies as reasons to approve new agents. In fact, the use of symptom scales and HRQOL measures in clinical trials has demonstrated that some new drugs have unexpected positive benefits for symptom control. As we learn more about the biologic bases of symptoms such as pain, nausea, vomiting, fatigue, and depression, the possibilities of symptom-focused interventions expand.

The Concept of Symptom Burden

The definition of “symptom” derives from the Greek word “symptoma,” which means “anything that has befallen one.” A more useful description is provided by Webster’s Third New International Dictionary (8) that defines a symptom as “the subjective evidence of disease or physical disturbance observed by a patient.” Implicit in this definition is the negative nature of symptoms and, most importantly, that symptoms are observations of the patient, the person experiencing the evidence of disease or physical disturbance. In contrast to “signs” of disease (such as fever or high blood pressure), symptoms can only be known through patient report.

It is frequently difficult for patients (and clinicians) to accurately ascertain the underlying basis of symptoms. Symptoms can be produced by the disease itself, or they can be produced by disease treatment, in which case they are often referred to as side effects or toxicities. Symptoms can also arise from comorbid medical conditions or acute injuries. Collectively, these sources of distress impose a “symptom burden” upon the patient that is a subjective counterpart of summary expressions of disease such as tumor or treatment burden (9). Symptom burden can be thought of as the sum of the severity and impact of symptoms reported by a significant proportion of patients with a given disease or treatment.

Symptom Burden and Cancer

Patients with cancer experience multiple symptoms that cause them significant distress and that impair function and rehabilitation. A 2003 report from the National Institutes of Health (10) stated that nearly 10 million people in the United States had a history of cancer and that another 1.3 million were expected to receive a cancer diagnosis in 2004.

Whereas many cancer-related symptoms are the result of disease, it is increasingly recognized that neuropathy, fatigue, sleep disturbance, cognitive dysfunction, and affective symptoms can also be caused by cancer treatment (11). Treatment-related symptoms may persist for weeks, months, or years and may worsen, even if the cancer itself improves. As patients survive cancer for increasingly longer periods of time, persistent residual treatment-related symptoms are becoming more prevalent and pose an increasing barrier to the resumption of predisease functioning. Treatment-related symptoms can directly affect survival if they become so severe that patients abandon important (and sometimes potentially curative) therapies or if they cause treatment delays (12). Posttreatment symptoms can also limit vocational activity and inhibit social recovery.

Symptom Clusters

Although the symptoms of cancer have most often been studied in isolation (e.g., studies of pain, studies of nausea and vomiting), anyone who has or has ever had cancer or who treats cancer patients knows that symptoms clearly occur together and can exacerbate one another. Multiple symptoms are additive in their impact upon patients with cancer (11,13–15) and significantly affect patient function. Various studies documenting the multiplicity of symptoms experienced by cancer patients (16–19) have shown that pain, fatigue, sleep disturbance, emotional distress, and poor appetite are almost universally found to be co-occurring. These nonspecific symptoms are not typically monitored as closely as toxicities in the clinical setting, and as a result, appropriate symptom management is often not addressed.

Two or more symptoms (for example, fatigue and appetite loss) that follow the same time course in response to disease or treatment have been designated as “symptom clusters” (13,20). An initial hypothesis of the causation of symptom clusters was that one symptom is liable to produce another (pain would naturally generate depression and sleep disturbance). However, many clinicians have observed that certain severe symptoms, such as pain, fatigue, difficulties with concentration, and sleep disturbance, appear together in sick patients across various diseases and treatments (16,20). It has therefore been postulated that this synchronicity is actually a symptom cluster with common underlying mechanisms. Although the mechanisms underlying the development of treatment-related symptom clusters in general are not well understood, there is a growing awareness that common biologic mechanisms (such as an inflammatory response produced by disease or treatment) may cause or contribute to some of these symptoms at the same time (11,21).

Not all symptoms increase or decrease together. Sets of symptoms or symptom clusters may have a different temporal pattern in relationship to treatment or disease progression (22), and improvement in one symptom (e.g., depression) may not be correlated with improvement in another symptom (e.g., fatigue) (23). These newer findings emphasize the importance of longitudinal assessment in potentially identifying the biologic basis of symptoms.

Assessing Multiple Symptoms

The definition of what constitutes a true symptom cluster is evolving with the development of multiple-symptom scales, several of which have been developed for use with cancer patients. An ideal multiple-symptom assessment tool should include symptoms that occur most frequently and are most distressing to patients. At the same time, the assessment should also be short, easy to understand, and applicable to both clinical and research settings. Given that symptoms have an adverse impact on function and activity, symptom scales should also assess the interference with different activities caused by these symptoms, as viewed from the patients’ perspective.
A recent systematic review of cancer symptom assessment instruments by Kirkova et al. (24) identified 21 of the tools as appropriate for clinical use and described in detail 14 of them that assayed more than five symptoms, rated by such criteria as comprehensiveness, psychometric properties, time to complete, and utility for decision making. Thirteen of the instruments assessed the patient’s perception of the impact of the symptoms, using scales that measure distress, bother, or interference with normal activities and life. Any of these scales can serve as a basis for measuring symptom burden so long as the patient’s evaluation of the salient symptoms interfere with functioning are represented.

Multiple-symptom inventories can be used to identify symptoms that are prevalent and distressing across different cancers and treatments. For example, the M. D. Anderson Symptom Inventory (MDASI), one of the instruments reviewed in the article of Kirkova et al. (24), is a brief measure of the severity and impact of cancer-related symptoms (16). The development of the MDASI was based on previous efforts assessing the severity and interference of single symptoms, including the Brief Pain Inventory and the Brief Fatigue Inventory (25, 26). The MDASI asks patients to rate the severity of 13 symptoms that are common in cancer patients once treatment begins: fatigue, sleep disturbance, pain, drowsiness, poor appetite, nausea, vomiting, shortness of breath, numbness, difficulty remembering, dry mouth, distress, and sadness. Patients rate each symptom’s presence and greatest severity in the previous 24 hours on an 11-point (0–10) scale, with 0 representing “not present” and 10 representing “as bad as you can imagine.” The MDASI also contains six items that describe to what degree the symptoms interfere with daily functioning, including work (including work outside the home and housework), relations with other people, and enjoyment of life. Each interference item is also rated on an 11-point scale with 0 signifying “does not interfere” and 10 signifying “completely interferes.”

Recognizing that symptoms do not occur in isolation and that patients typically experience multiple symptoms caused by disease or treatment, a measure of symptom burden might be a summative indicator of 1) the severity of the symptoms most associated with a disease or treatment and 2) the patient’s perception of the impact of these same symptoms on daily living. In contrast, HRQOL is best viewed as a subjective evaluation of life as a whole (2). In the conceptual model of HRQOL, the patient’s perception of the impact of symptoms goes beyond the reporting of symptom severity into the more abstract concepts included in the meaning of HRQOL. The model does, however, limit questioning of impact to the patient’s impressions of the impact of specific symptoms or symptom clusters.

The comprehensive nature of HRQOL is both one of its attractions and one of its difficulties as a PRO measure. Intuitively, a significant reduction in symptoms should bring improvement in other aspects of HRQOL, but does this necessarily happen? Jatoi et al. (28) commented that, if a symptom is reduced but a benefit is not demonstrated in more generic measures of HRQOL, the treatment should still be provided. A review of symptom management trials conducted under the auspices of the Community Clinical Oncology Program (7) concluded that the value added by including broader HRQOL measurement in symptom trials has yet to be demonstrated. The authors pointed out that HRQOL is often presented as a secondary endpoint and that a conceptual connection between symptom reduction and changes in HRQOL is often not delineated.

### Why Symptom Burden Outcomes May Be Sufficient

Reducing symptom burden, with or without a demonstration of positive impact on generic HRQOL, whether in trials with a symptom endpoint or trials comparing otherwise equivalent curative therapies, may well provide the evidence needed to make optimal choices, both in treatment and in defining clinical trial endpoints. We therefore propose that the measurement of symptom burden may be a sufficient outcome in many clinical trials, giving health providers, patients, and regulators enough information to make decisions about whether to use, approve, or fund a specific treatment.

Studies have shown that patients may report their HRQOL as improved or unchanged despite changes or deterioration in health. In such situations, HRQOL measures are liable to yield discrepant

### Table 1. Percent moderate-to-severe for each symptom (arranged in decreasing magnitude) by disease*

<table>
<thead>
<tr>
<th>Symptom</th>
<th>Breast cancer</th>
<th>Lymphoma</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fatigue</td>
<td>51.10</td>
<td>67.10</td>
</tr>
<tr>
<td>Sleep disturbance</td>
<td>36.30</td>
<td>41.90</td>
</tr>
<tr>
<td>Drowsiness</td>
<td>31.90</td>
<td>39.20</td>
</tr>
<tr>
<td>Distress</td>
<td>30.80</td>
<td>36.50</td>
</tr>
<tr>
<td>Pain</td>
<td>29.20</td>
<td>36.50</td>
</tr>
<tr>
<td>Difficulty remembering</td>
<td>28.30</td>
<td>31.10</td>
</tr>
<tr>
<td>Sadness</td>
<td>26.70</td>
<td>30.60</td>
</tr>
<tr>
<td>Dry mouth</td>
<td>25.80</td>
<td>29.70</td>
</tr>
<tr>
<td>Numbness</td>
<td>23.90</td>
<td>28.40</td>
</tr>
<tr>
<td>Lack of appetite</td>
<td>21.10</td>
<td>27.10</td>
</tr>
<tr>
<td>Nausea</td>
<td>12.20</td>
<td>23.00</td>
</tr>
<tr>
<td>Shortness of breath</td>
<td>7.80</td>
<td>14.90</td>
</tr>
<tr>
<td>Vomiting</td>
<td>13.20</td>
<td>24.70</td>
</tr>
</tbody>
</table>

* In the case of three symptoms (fatigue, lack of appetite, and shortness of breath), the proportion of patients with symptoms at the severe level is higher in the lymphoma group.
† P < .038.
‡ P < .029.
§ P < .039.
results compared with symptom burden measures. Studies of patients who have undergone bone marrow transplantation have shown patients rating their quality of life as above average despite the persistence of significant physical and psychologic symptoms. For example, Bush et al. (29) conducted a descriptive study of quality of life, psychologic distress, demands of long-term recovery, and health perceptions of 125 survivors of bone marrow transplantation. They were no different from individuals sampled from the general population with regard to their responses on HRQOL measures. However, 10 or more years after transplantation, long-term survivors continued to experience a moderate incidence of lingering complications, including emotional and sexual dysfunction, fatigue, eye problems, sleep disturbance, general pain, and cognitive dysfunction.

Several recent studies have suggested that symptom measures, in contrast to global HRQOL measures, may provide information that is more indicative of treatment differences. For example, in a clinical trial comparing raloxifene with tamoxifen (30), no significant differences were reported between the two groups in PROs for physical health and mental health, which are generic components of HRQOL, even though the tamoxifen group reported better sexual function and the two groups differed significantly in the clusters of symptoms they reported. In another study of adjuvant therapy for breast cancer, Fallowfield et al. (31) found that 2 years of treatment with anastrozole, tamoxifen, or a combination of the two had a similar overall impact on HRQOL, showing gradual improvement over time, but that different symptom profiles were associated with the two agents. The researchers concluded that the different symptoms experienced by patients in separate arms of the trial may assist in decision making about treatment and supportive care needs.

In a recent study of patients’ perceptions of the side effects of treatment for prostate cancer, Korfage et al. (32) noted a relative lack of change in HRQOL scores over time despite increases in sexual, urinary, and bowel symptoms. The authors suggested that persistently high generic HRQOL scores cannot be interpreted as “these men are doing just fine” and challenged the conclusion of Krahn et al. (33) that, since HRQOL remains unchanged, the side effects of early prostate cancer treatment need not be of concern. Bottomley et al. (34) reviewed 24 clinical trials in the EORTC that included HRQOL measures. The authors identified 13 trials with differences in PROs between trial arms. Of these, differences in symptom report (pain, fatigue, neuropathy) were reported in nine. Differences between arms in global quality of health measures were reported in four studies, and nonsymptom components of HRQOL were reported in one additional study. The contribution of changes in symptoms in the differences found in these studies was not separately studied.

A recent EORTC clinical trial compared first-line chemotherapies for breast cancer (doxorubicin and cyclophosphamide versus doxorubicin and paclitaxel) (35). While no differences were seen in a global HRQOL measure, significant increases in treatment-related symptoms, including nausea and vomiting, were reported in both groups. Fatigue also increased, but it remained high in the paclitaxel arm only. However, pain was reduced in both groups, consistent with a treatment effect. The authors concluded that, although there were clinically meaningful increases in treatment-related symptoms, these effects did not adversely influence global HRQOL scores, which were stable across treatments. This study illustrates the potential differences between symptom measures and more generic HRQOL outcome measures. It may well be that reporting expected increases in symptom burden associated with these treatments (together with the possible reduction in pain) might be more informative to patients and their clinicians than reporting that HRQOL would not be expected to change.

**Future Directions for the Development and Use of Multiple-Symptom Measures**

Methods for assessing single symptoms and their impact are relatively well established, and there is beginning to be consensus about how single-symptom–focused clinical trials should be conducted (for example, see the summary of the work of the IMMPACT [Initiative on Methods, Measurement, and Pain Assessment in Clinical Trials] group on clinical trials in pain (36)). Methods for assessing multiple symptoms and their impact, which we have called symptom burden, are less developed, although we now have several multiple-symptom tools that can be used. Since both the American Society of Clinical Oncology and the European Society of Medical Oncology have endorsed pain management, supportive care, palliative care, and survivorship planning as key aspects of quality cancer care, the importance of impeccable symptom measurement cannot be overstated (37).

Optimal use of a multiple-symptom measure entails recognizing that there may be highly frequent and distressing symptoms unique to a specific disease, disease stage, or treatment. For example, diarrhea is a common sequela to several treatments, especially for gastrointestinal cancers, whereas constipation is common at a later stage, when opioid analgesics are frequently used (16). If symptoms or symptom burden is chosen as an outcome variable of importance, sufficient identification of the most relevant symptoms for the specific disease or treatment to be studied must be incorporated into the study design.

In one approach to this problem, a recent study (38) used the EORTC QLQ-C30 to examine the preferences of patients in different disease groups for functional domains and symptoms. For the group as a whole, role, cognitive, and social functioning; fatigue; nausea and vomiting; pain; and appetite loss; diarrhea; and financial difficulties were the most important effects to avoid, whereas physical and emotional functioning, dyspnea, constipation, and insomnia were seen as less important. The rankings varied, however, when responses were categorized by disease group. The four effects that patients with breast cancer most wished to avoid were nausea and vomiting, pain, and decreases in emotional and role functioning, whereas for patients with non–small-cell lung cancer, dyspnea was the fourth most important effect to avoid. Patients with colorectal cancer listed nausea and vomiting, diarrhea, pain, and decreases in role functioning. As chosen by the patients, the effects to avoid were consistent with well-recognized symptoms and treatment side effects that would likely be experienced by patients with these different cancers.

New therapies bring with them new treatment-related symptoms. For example, rash, a symptom not seen in older chemotherapy treatments, is becoming recognized as a disturbing symptom among newer, targeted therapies (39), yet few, if any, of the
current multiple-symptom scales include rash as an item. The simplest solution to this problem would be to add a missing symptom item to a previously validated symptom assessment instrument, but the need to demonstrate the sensitivity and psychometric characteristics of these new items is a topic of debate (1). Does adding a single or a few additional symptoms to existing scales require complete revalidation? Probably not, but components of validity and sensitivity need to be demonstrated for these items before their incorporation into an already validated multiple-symptom scale.

It should be noted that, in many instances, making a decision about cancer treatment options requires more than characterizing symptom severity, symptom interference, or symptom distress, the information typically provided by a PRO measure. For example, if treatment choices would yield differences in post-treatment appearance, the patient’s perceptions of the impact of these differences are potentially important outcome measures in selecting the appropriate therapy. Data on long-term restoration of function (social, vocational, and physical) are also extremely important when making treatment decisions, but these data are rarely available from PROs obtained during a clinical trial.

In summary, multiple-symptom and single-symptom scales, if conceptually justified by the trial design and intent of the therapy, may provide outcome data sufficient to make decisions about the value of a therapy or to allow judgment about the relative value of one therapy contrasted with another. Many long-term outcomes that fit within the HRQOL domain (social and role function, for instance) will not be known until long after the trial is completed. Several of the domains of HRQOL (such as role and social functioning and concern with financial problems) that are more distal to the treatment process (and less liable to be influenced by whatever treatment is being evaluated) need conceptual support before they are included as trial outcomes.

References

(31) Fallowfield L, Cella D, Cuzick J, Francis S, Locker G, Howell A. Quality of life of postmenopausal women in the Arimidex, Tamoxifen, Alone or in...


