Patient Symptoms and Clinician Toxicity Ratings: Both Have a Role in Cancer Care

Cardyn Gotay

Basch et al. (1) contribute to a growing literature on the concordance—or lack thereof—between ratings made by cancer patients and proxies. This question is not new: A classic 1992 review by Sprangers and Aaronson (2) included 35 articles that explored the relationship between clinician and cancer patient assessments of quality of life (including symptoms). Interestingly, one of the summary conclusions of this review was that health-care providers underestimate the pain intensity reported by patients, a conclusion that is echoed in the study by Basch et al.

However, much has changed in this field since 1992, and the study by Basch et al. reflects these clinical and methodological advances, including the increasing acceptance of obtaining patient ratings in the therapeutic environment, and the use of standardized and previously validated questionnaires. This study is particularly noteworthy for its collection of comparable data from providers and patients at the same time points, the impressive response rates for both patients and providers, and the access to full information about specific outcomes: patient use of cancer care services (i.e., emergency room [ER] visits) and vital status. Furthermore, the setting of primarily advanced small cell lung cancer with a 41% mortality rate during the study period provided an opportunity to examine these outcomes.

It is often not acknowledged that there is error, and bias, in ratings from any source, be it patient or provider. Patient ratings are influenced by factors such as the individual’s motivation, interpretation, expectations, and personality (3). For example, there is evidence to support the importance of dispositional optimism: optimistic patients experience less pain (4). Sociocultural factors also have an influence on whether a particular symptom gives rise to patient-rated distress and whether the patient wishes to convey this to others. For example, both the meaning and the expression of pain are affected by cultural factors, such that Chinese are reluctant to report pain to others (5). Thus, symptom ratings in patients are affected not only by the disease and treatment but also by other characteristics the patient brings to the situation. Furthermore, cancer patients are limited by their own experience; if they have never experienced severe pain previously and are asked how severe their pain is, how are they to know how much more severe it may become? This study implies that patients reached their “rating ceiling,” or greatest possible pain rating earlier than clinicians, perhaps for this very reason.

Clinicians have the advantage, if it can be called that, of knowing that even if patients report high levels of pain early on, their symptoms may get worse. This knowledge may lead them to have a higher ceiling for their ratings because they are able to distinguish between higher levels of pain more precisely based on their previous experiences. Who is more accurate: the patient who says “This pain is as bad as it can be—I rate it 10 out of 10,” or the clinician who thinks, “It’s going to get worse. She’s probably only at an 8?”

The more important question than which source is right is “how can this information be used?” The study by Basch et al. provides a compelling answer to this question: Clinician ratings, when they assess a patient’s symptom at a common terminology criteria for adverse events (CTCAE) threshold severity level, were strongly predictive of subsequent patient use of the ER and death. I disagree with the authors that the strong correlation between the ratings and outcomes implies that the information is redundant and that clinician-reported symptoms add only limited value. Instead, it would seem that such ratings provide critical information for care planning—and for possible prevention of ER visits. The article did not provide information on reasons for ER admissions, or patterns of care in these patients, but if additional prospective monitoring or hospice referral were initiated based on the clinician ratings, perhaps ER visit rates could be reduced.

How can the patient information be used? The authors’ suggestion that patient symptoms reflect “daily health status” is not convincing because the EQ-5D measure was administered at the same time as the symptom measures and includes items that directly map the individual symptom scores. It is not surprising that the EQ-5D scores are closely linked with patient symptom scores because they reflect the same perspectives and sources of error and bias. The global score, because it was only collected at baseline, is difficult to interpret in terms of changing symptom and toxicity ratings.

Surprisingly, the authors do not note that patient severe pain ratings were in fact linked with a statistically significantly higher risk of death and are marginally statistically significant (P = .06) for ER admission. As the authors state, patient ratings of other symptoms were not linked with these outcomes. Why might pain ratings be more similar for clinicians and patients? The answer to this question, which has also been noted by others (6), may relate to one of the most important potential sources of clinician error: the lack of a systematic approach to gathering data to assess toxicity. Without documentation of how clinicians made judgments of toxicity ratings for the symptoms measured in this study, we cannot be sure of what information they used. However, it seems likely that for the most subjective symptom—pain—they likely asked the patient, and the patient’s own appraisal contributed to the clinician’s rating, thus explaining the higher concordance in pain outcomes.

Patient ratings are critical in cancer care to understand the impact of the disease and treatment on the individual and to develop
appropriate supportive measures. When it comes to using this information for labeling purposes, the suggestion in the article—that patients act as consumers and sort out the relative usefulness of professional vs patient data—seems to be asking a bit much of patients. However, presentation of both patient and provider data is justified, with the findings described in terms of the specific ways the questions were asked, rather than global statements about “adverse events.” For (hypothetical) example, the label might state: “While most patients experienced nausea, the majority (50%) rated it as mild, although 25% said it was ‘very severe.’ Clinicians found that 20% of patients had such severe nausea that they couldn’t eat or drink normally and required medical intervention” (which is the CTCAE grade 3 nausea criterion). Patients themselves need to be involved in developing such descriptions, so that the information they need and want is available and understandable.

References