The Biology Behind

Is There a Role for the Assessment of Health-related Quality of Life in the Clinical Evaluation of Novel Cytostatic Agents?


Linda J. Patrick-Miller
Department of Psychiatry, University of Medicine and Dentistry of New Jersey-Robert Wood Johnson Medical School, The Cancer Institute of New Jersey, New Brunswick, NJ 08903

Introduction
In this issue of Clinical Cancer Research, LoRusso et al. (1) report on the longitudinal assessment of HRQOL(2) that was performed in two multicenter Phase I trials of ZD1839 (“Iressa”) in patients with NSCLC, head and neck, prostate, ovarian, or colorectal cancer. This report is novel, not because the safety and tolerability of Iressa, or the HRQOL of patients undergoing treatment with Iressa for solid tumors is reported (2–4), but because HRQOL assessed in the context of these Phase I trials is reported. The rationale for this exercise is the evaluation of the potential for HRQOL assessment to serve as a measure of treatment efficacy that would facilitate the quantification of clinical benefit for this novel cytostatic anticancer agent. To address the question, “is there a role for the assessment of health-related quality of life in the clinical evaluation of novel cytostatic agents?” we first review the traditional end points used in the clinical evaluation of novel cytotoxic agents, the emerging trend to identify new end points consistent with points used in the clinical evaluation of novel cytotoxic agents, and the historical contribution of HRQOL assessment to anticancer treatment outcome research.

Biological End Points in the Evaluation of Chemotherapeutic Agents

Historical: Cytotoxic Agents. Prior to the last decade, the biological mechanism of most potential novel chemotherapeutic agents was cytotoxicity, and the hypothesized effect was tumor shrinkage. Clinical trials designed to evaluate novel agents were based on this assumption, and outcomes were selected accordingly. Historically, the clinical evaluation of a new chemotherapy agent has involved a three-phase process: (a) a Phase I trial to determine appropriate dose for additional testing; (b) a Phase II trial to determine the biological activity (efficacy) of the agent; and (c) a Phase III trial to determine the relative strength of the biological activity of the novel agent compared with a standard agent of known effectiveness.

For Phase I trials, toxicity was the end point. To assess toxicity, symptoms or signs that occur in individual patients while on trial are each rated for likelihood of association with the agent (rather than the disease), and severity. Ratings are made according to criteria established for each sign/symptom by oversight agencies.3 In an inductive process, toxicities of an a priori defined severity (i.e., DLT = CTC ≥ 3) are aggregated across patients to approximate group results. An acceptable proportion of this group that will experience DLTs is also defined a priori (e.g., 30%). The dose level associated with this threshold is identified as the MTD. This MTD is exported to the Phase II trial, as the dose at which the novel agent will be evaluated for biological effect and comprehensive assessment of toxicity. Thus, the quality of this determination process and the accuracy of the output are critical to the additional evaluation and development of the novel agent; if the MTD is too low, the biological activity may be underestimated; conversely, if the MTD is too high, the toxicity may be overrated (5).

In Phase II trials, the biological effect of cytotoxic agents is evaluated on the basis of the changes that occur in the size of the targeted tumors while undergoing treatment with the novel agent. Measurements are taken of each targeted lesion at baseline and designated intervals. For each patient, at each time point, the measurements are summed across lesions. Percentage of changes in those summed measurements, as compared with baseline, are calculated (e.g., 30% decrease), and categorized using established criteria by extent (e.g., partial) of objective tumor response (6). The best objective tumor response for each patient is selected, and categorized into established criteria of overall response (e.g., Response Evaluation Criterion Solid Tumors).4 In an inductive process, the proportion of the patients who have overall responses at each designated level (e.g., stable disease, or partial or complete response) is used to estimate the expected level of biological activity of the agent in the corresponding population. A proportion of patients with a given overall response level may be used to define a priori a minimal

---

1 To whom requests for reprints should be addressed, at Division of Behavioral Sciences, The Cancer Institute of New Jersey, 195 Little Albany Street, 3039, New Brunswick, NJ 08903. Phone: (732) 235-8078; Fax: (732) 235-7493; E-mail: patmiller@umdnj.edu.

2 The abbreviations used are: HRQOL, health-related quality of life; NSCLC, non-small cell lung cancer; CTC, common toxicity criteria; DLT, dose-limiting toxicity; MTD, maximum tolerated dose; FDA, Food and Drug Administration; CMC, clinically meaningful change.


biological effect (e.g., 20% of patients). Novel agents that reach the minimal biological activity level move on to Phase III evaluation.

**Evolving: Cytostatic Agents.** In recent years, the focus of discovery of potential anticancer agents has shifted toward the development of drugs directed at specific molecular targets. The biological mechanisms of these agents, such as ZD1839 (Iressa), are cytostatic rather than cytotoxic. The hypothesized effects of cytostatic agents may include, for example, tumor growth inhibition, rather than tumor shrinkage, and these effects may be achieved with less toxicity to normal tissue (7, 8). This shift has prompted concern that Phase I and II trial designs based on the assumption of a dose-dependent tumor shrinkage effect that is correlated with clinical benefit (as well as toxicity), which historically have been used to effectively evaluate novel cytotoxic agents, may not provide the optimal opportunity for evaluation of potential cytostatic agents (7–10). These concerns have prompted suggestions that alternative trial designs be adopted and additional end points incorporated into the evaluation of novel cytostatic agents.

For Phase I trials, it has been suggested that biological measures consistent with the targeted activity of the agents that have been validated in preclinical studies (e.g., inhibition of epidermal growth factor receptor activation) may provide a means for defining the biologically effective dose, that could be used as a complementary or alternative end point to DLT (8, 10). Similarly, for Phase II trials, the proportion of patients exhibiting a predefined percentage of change in tumor size may not provide an accurate population estimate of the efficacy of the cytostatic agent (7–9). Additional suggestions have included the proportion of patients with early progression or progression-free survival time and changes in tumor markers (7, 8).

Caution has been urged in adopting end points other than increased survival time, the “gold standard” of novel drug evaluation, or tumor response rate, its surrogate (8). However, a recent review of drugs approved from 1990 through 2002 by the FDA Division of Oncology Drug Products in the Center for Drug Evaluation and Research suggests that this caution may be overstated (11). Although FDA approval of oncology drugs was previously determined solely on the basis of tumor response rate, this position was revised in the mid-eighties when it was noted that the toxicities of these agents did not necessarily substantiate the presumed relationships among response rate, survival time, and “clinical benefit.” The revised FDA position requires that evaluation studies include a quantitative assessment of the efficacy of the agent to either prolong or improve life (12). This opened the door for novel agents to be approved based on their capacity to improve HRQOL.

**HRQOL Assessment in Clinical Trials**

**Historical: Phase II-III Trials.** Interest in the quality, as well as the duration of life has been an implicit consideration in the delivery of medical care since at least the time of Hippocrates, who bid us “do no harm” (13). However, interest in the explicit scientific investigation of HRQOL was not notable until the mid-seventies (14). Over the next decade, as the FDA added efficacy of an anticancer agent in improving life as an acceptable outcome for approval (12), the first studies assessing HRQOL in conjunction with cooperative group Phase II-III trials were conducted (15, 16).

HRQOL, not unlike clinical benefit, is a broad, multidimensional construct. Definitions vary, but there is general agreement that HRQOL is a subjective sense of the psychological, emotional, and physical functioning of a person that includes disease- and treatment-associated symptoms (17). HRQOL assessments have been used in many ways, such as the prediction of survival time, and enrollment in Phase I trials (18, 19). However, the predominant application of HRQOL research has been the evaluation of treatment outcomes in randomized clinical trials. In fact, many oncology trial groups now have policies requiring that HRQOL outcomes be considered in the design of all Phase III trials, and if not included, this decision justified (16, 20).

Typically, HRQOL data are collected longitudinally in conjunction with biological data from all of the patients throughout their participation in a Phase III trial. The data are then applied in a deductive process to estimate the clinical significance of the findings of the study sample for an individual patient. However, the natural history of the disease, the expected impact of the novel agent, and the efficacy of available treatments determine the design of the clinical studies, and the nature of the outcomes data to be collected. For example, when the disease is symptomatic and there is no known effective treatment (pancreatic cancer), the expected impact of the novel agent (gemcitabine) may include symptom palliation (a surrogate for a “better life”), as well as survival rate (a surrogate for a “prolonged life”). Patient-reported assessments of disease-related symptoms were included in the Phase II trial of gemcitabine. The output of the Phase II trial was the definition of a composite measure of clinical benefit that served as the primary end point in Phase III testing and the basis of FDA approval (7, 11, 21, 23).

To facilitate its inclusion in the evaluation of clinical benefit for novel agents, researchers have applied the concept of CMC to HRQOL assessment. The concept applies to both cross-sectional differences (e.g., between treatment arms) and longitudinal differences (e.g., pre-posttreatment). Two complementary methodologies are used jointly to empirically define CMC. Anchor-based methods describe differences or change in the index measure over time (e.g., HRQOL) in terms of an external criterion identified as clinically relevant in the study context (e.g., tumor response; Refs. 23, 24). For example, NSCLC patients who responded to treatment had (statistically) significantly lower scores on a HRQOL measure after two treatment cycles than nonresponders, and their scores dropped (statistically) significantly more during that time, providing an estimate of the CMC in HRQOL associated with tumor response in this study population (25). Distribution-based methods evaluate the magnitude of change in the index measure (e.g., HRQOL) in conjunction with an indicator of the variability of the measure (e.g., SE of the measurement) within a given study to generate an effect size that constitutes a CMC (20). Statistical assessment of the consistency of the products of each of these methods provides a measure of confidence in the calculated CMC (23, 25).
Evolving: Potential Contribution of HRQOL Assessment to Phase I Trials. The development of empirically derived indices of CMC in HRQOL provides an opportunity to use these measures in the development of end points for the evaluation of novel cytostatic agents, to supplement or replace traditional outcomes that may be less specific to the imputed mechanisms of action of these agents, and, therefore, less sensitive to their effectiveness. Particularly promising is the opportunity to more clearly define and operationalize the concept of clinical benefit. However, like all tools, indices of CMC in HRQOL are no more effective than the rigor with which they are applied. Detecting meaningful change depends on the appropriateness of study design to the context, the care of the data collection, and the quality of the data analysis.

Missing data has been the bane of HRQOL assessment since its earliest inclusion in clinical trials (15, 26, 27). However, evaluation of missing data has demonstrated that although both institutional and patient factors contribute to avoidable failure to collect HRQOL data, patient refusal is minimal (7%), and institutional factors are the greatest contributors to baseline noncompliance (14, 20, 28). In missing data, as in cancer, prevention remains the best treatment (14), and compliance rates have improved considerably since the earliest trials, averaging 85% in the 1990s (28). Missing data cannot be assumed to be random (14), and must therefore be examined both cross-sectionally and longitudinally for patterns suggesting the need for specific data analytic techniques (14, 28–30). Missing baseline data, as the essential component for the calculation of any treatment associated change, is irreplaceable in any trial (14, 19).

Although some design and analysis concerns are universal in the assessment of change in HRQOL, other factors are either unique to, or exaggerated in, the Phase I setting. For example a greater proportion of “nonignorable” missing data may result from disease progression or death in Phase I than in later-phase trials (31). Other potential biases involving evaluation of change over time include the overgeneralization of population-specific criterion-based estimates of CMC to Phase I study samples that tend to be more heterogeneous than later-phase trial populations in, for example, disease and performance status (24). The small and heterogeneous sample affects statistical significance, as well as, clinical significance. In a heterogeneous sample, the effect size of a change in HRQOL may be small compared with the between-person SD, and, therefore, interpreted as trivial, whereas same effect size in a more homogeneous population would appear more (statistically and clinically) significant as a result of the smaller SD (23, 32).

Although not trivial, these details of data analysis and interpretation side-step a broader issue involved in bringing to bear the deductive process of using population-derived CMC to estimate the impact of treatment for individuals to the Phase I context, where individual assessments are applied inductively to estimate population effects. The inductive application of HRQOL data are not inconceivable, but are also not yet well-accepted (24).

What function then might assessment of HRQOL serve in the context of the Phase I trials? Two areas hold particular promise. The first is the potential to better inform the output of Phase I trials, the dose at which biological efficacy and toxicity testing will occur in Phase II evaluation. The current determinant of this critical output, DLT, has been criticized for its lack of sensitivity and specificity for cytostatic agents (7–10). Healthcare provider estimates consistently assign lower severity ratings to patient somatic symptoms than do patients (33–35), and CTC for symptoms that patients may report as most concerning can be range-restricted (e.g., alopecia) or skewed (e.g., macular/papular rash; Ref. 1). Thus, patient-reported assessments of treatment-associated symptoms have the potential to provide a sensitive and specific evaluation of treatment-associated symptoms to supplement the physician ratings using the CTC. This symptom assessment could facilitate an estimate of the MTD, as well as serve as the pilot of the instrument to be used in the more comprehensive assessment of tolerability in the Phase II trials.

The second particularly promising opportunity for the assessment of HRQOL in Phase I trials is the potential to begin to more specifically define and operationalize the concept of clinical benefit. It has been suggested that for response rate to truly serve as a surrogate for clinical benefit, it would need to account for response duration, and treatment-associated and disease-associated symptoms (11). HRQOL assessment could contribute to these symptom-based assessments. More importantly, preliminary assessments of the relationships between biological measures consistent with the targeted activity of novel agents could be used to define criteria for the establishment of CMC for evaluation in Phase III studies.

Conclusion. The investigators of the Phase I Iressa trials are to be applauded for their efforts to raise the utilization of HRQOL outcome data in the evaluation of novel anticancer agents to the next level through this assessment of its feasibility and utility in this context. The critical question to be asked as we examine these exploratory efforts is what contribution might HRQOL data make to the evaluation of novel anticancer agents in early clinical testing. Likewise, once the purpose of assessing HRQOL has been defined, how might we draw on the considerable science and methodology that has been brought to the study of HRQOL over the past two decades to investigate this potential with at least the same rigor as we apply to other aspects of the evaluation of novel anticancer agents?

It has been suggested that methodological shortcomings of clinical trials that include assessment of HRQOL have limited the contribution of this data to the evaluation of anticancer drugs (11, 36, 37). A meta-analysis of 25 randomized clinical trials conducted between 1980 and 2001 that included HRQOL assessment lends support to this opinion. The inclusion in this analysis of the very earliest RCTs that assessed HRQOL surely explains some reported shortcomings for which there have been considerable methodological advancements in the past two decades, such as the treatment of missing data (14, 28–30). However, much more striking is the proportion of studies that failed to adhere, not only to published guidelines for the conduct of HRQOL research (16, 20, 38–42), but to the most basic tenets of scientific investigation. For example, only 12% of studies analyzed included clearly stated hypotheses regarding HRQOL (36)!

HRQOL research, like other aspects of clinical oncology research, is a less-than-perfect but vigorously evolving science. Just as there is no one currently agreed upon “gold standard” for
the assessment of the biological effect of cytostatic agents, there is not a single, universally agreed upon HRQOL methodology to inform their psychological, social, and physical impact. This should not preclude the rigorous application of the considerable science and methodology that has been brought to the study of HRQOL (43, 44). It is through these cross-disciplinary collaborations that multidimensional outcomes research will rise to meet the evolving needs of the clinical investigation of novel anticancer agents, cytostatic, as well as cytotoxic.

References


Linda J. Patrick-Miller


Updated version  Access the most recent version of this article at: http://clincancerres.aacrjournals.org/content/9/6/1990

Cited articles  This article cites 35 articles, 7 of which you can access for free at: http://clincancerres.aacrjournals.org/content/9/6/1990.full#ref-list-1

E-mail alerts  Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions  To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions  To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.