

## Reflections on Findings of the Cancer Outcomes Measurement Working Group: Moving to the Next Phase

*Carolyn C. Gotay, Joseph Lipscomb, Claire F. Snyder*

---

**The Cancer Outcomes Measurement Working Group (COMWG) was a National Cancer Institute working group of 35 experts convened to examine the state of the science and identify future research priorities for outcomes assessment in cancer. The COMWG focused on three outcomes (health-related quality of life [HRQOL]), patient needs and satisfaction, and economic burden) in four cancers (breast, colorectal, lung, and prostate) across the continuum of care (prevention and screening, treatment, survivorship, and end of life). The majority of the research to date has focused on HRQOL assessment, which has been shown to be feasible in a research context, using questionnaires that meet established criteria for reliability and validity. The quality and quantity of HRQOL research has increased markedly in recent years, and additional methodological developments—particularly the application of item response theory to improve precision, efficiency, and comparability in measurement—hold considerable promise. Research is needed to develop and test predictive models of HRQOL and to establish the added value of including HRQOL assessment in clinical trials. [J Natl Cancer Inst 2005;97:1568–74]**

---

Persons who undergo cancer screening, diagnosis, treatment, or continuing care are in a position to contribute unique perspectives on the quality of cancer care. These perspectives can provide important information to the decision makers who provide, pay for, and regulate cancer care. However, measuring so-called “patient-reported outcomes” poses challenges that are different from assessing biomedical endpoints such as mortality and survival. Outcomes that rely on patient reports such as health-related quality of life (HRQOL) are inherently subjective; are not verifiable by objective indicators, such as physiological values; and are not generally part of standard clinical information systems.

As part of its Quality of Cancer Care Initiative, the U.S. National Cancer Institute (NCI) convened the Cancer Outcomes Measurement Working Group (COMWG) in 2001 to assess the current state of the science and to provide recommendations for future research to improve the assessment of these patient-reported outcomes. The COMWG was made up of 35 scientists from the United States, Canada, and the Netherlands; its members included medical oncologists, psychologists, economists, biostatisticians, nurses, health services researchers, and cancer survivors, with members from government agencies, universities, and cancer centers (see Appendix).

The COMWG focused on three primary outcomes—HRQOL, patient needs and satisfaction, and economic burden—in the four

cancers that affect the largest numbers of individuals in the United States: breast, colorectal, lung, and prostate. The group considered outcomes measurement across the entire spectrum of cancer prevention, detection, treatment, survivorship, and end-of-life care. COMWG members provided critical review and discussion of current literature in prespecified topics and identified areas in which more work is needed. The specific research questions that COMWG members were asked to address were as follows: What are the current best practices and recommendations for the future in assessing the three outcomes (HRQOL, satisfaction, economics) in the four cancers (breast, colorectal, lung, prostate) across the continuum of care? What are the definitions and conceptual models for outcomes measurement? What are the methodologic issues and challenges (psychometric, statistical, study design, and execution)? What are different perspectives on and uses of cancer outcomes data?

Methods used to address these questions included extensive searches of the research literature, searches of published and unpublished reports of government agencies, new analyses of existing datasets, and interviews with focus groups. Because the COMWG was not a federal advisory committee or consensus panel, its findings reflect only the individual judgments and perspectives of its members. Full and detailed accounts of the analysis of specific topics addressed by COMWG, including information about the specific methods used, have been published elsewhere (1).

In this commentary, we highlight some of the overarching themes that emerged from our analysis of the COMWG effort. Our commentary should not be construed as reflecting a consensus of the COMWG membership. Most of our discussion focuses on HRQOL because this area has received more attention in the research literature than the other outcomes; however, we include the other outcomes when appropriate. We focus on 1) the quality of the tools currently available to measure HRQOL in cancer research, 2) methodologic accomplishments and promising directions, and 3) use of HRQOL measures and interpretation of findings. We conclude by identifying research priorities for moving the field forward.

---

*Affiliations of authors:* Cancer Research Center of Hawaii, University of Hawaii, Honolulu, HI (CCG); Rollins School of Public Health, Emory University, Atlanta, GA (JL); School of Medicine, Johns Hopkins University, Baltimore, MD (CFS).

*Correspondence to:* Carolyn C. Gotay, PhD, Cancer Research Center of Hawaii, University of Hawaii, 1236 Lauhala Street, Honolulu, HI 96813 (e-mail: cgotay@crch.hawaii.edu).

See “Notes” following “References.”

DOI: 10.1093/jnci/dji337

© The Author 2005. Published by Oxford University Press. All rights reserved. For Permissions, please e-mail: journals.permissions@oxfordjournals.org.

## THE QUALITY OF CURRENT TOOLS USED TO MEASURE HRQOL OUTCOMES IN CANCER

COMWG members evaluated the quality of current outcomes measures with the instrument review criteria recommended by the Medical Outcomes Trust (MOT) (2). In accordance with the MOT paradigm, the performance of each instrument was assessed with respect to its conceptual and measurement model, reliability, validity, responsiveness, interpretability, respondent and administrative burden, alternative forms (i.e., modes of administration), and cultural and language adaptations (including instrument translation). The MOT paradigm was useful for comparing HRQOL instrument performance in many different areas, including cancer sites [breast (3), colorectal (4), lung (5), and prostate (6)], end-of-life care (7), caregiver impact (8), generic measures (9), patient perspectives on care (10), needs assessment (11), and survivorship (12).

Application of the MOT framework to the literature about use of HRQOL instruments gave rise to the following inferences:

**1) Despite skepticism by some critics, assessing HRQOL in a research context is feasible.** Patients do not resent being asked about personal aspects of their lives, and they do not object to using numerical, somewhat abstract scales to indicate their responses. HRQOL assessment can be incorporated in clinical protocols without placing undue burdens on either patients or clinicians. HRQOL assessment can be accomplished in the context of clinical trials, the application in which the largest amount of research has focused to date; in observational studies of treatment outcomes (13); and in descriptive studies of the impact of cancer in selected populations (14).

**2) It is widely (although not universally) agreed that HRQOL is an unobservable, or “latent,” construct that acquires operational meaning only through an appropriately specified and estimated psychometric measurement model.** A key aspect of HRQOL assessments (as well as patient reports, evaluations of care, and needs assessments) is that these outcomes reflect the patient’s perspective. There is also general agreement that these outcomes should be regarded as multidimensional constructs. In particular, HRQOL nearly always includes measures of physical and mental/emotional health, although there is less consensus about whether additional domains, such as social functioning, spiritual well-being, symptoms, and role functioning, are essential to HRQOL. In addition, specific domains may be particularly relevant in a limited set of circumstances (e.g., sexual functioning in prostate cancer, fertility concerns in survivorship). Measures of patient satisfaction or care experiences frequently include domains such as provider communication and access. The domains—that is, the general categories—of patient needs tend to be much less clearly specified than those of HRQOL. The more common measurement strategy is therefore the assessment of specific and well-defined needs, as opposed to broader domains.

**3) Numerous HRQOL questionnaires are currently available that have been developed with careful consideration of most of the MOT criteria (2).** Some of the most common approaches to HRQOL assessment include (a) generic measures, which are designed to be used in any disease group or a general population; (b) general cancer measures, which have been developed for use across cancer types; (c) cancer-specific measures, which are designed to be used in specific cancers; (d) general cancer core questionnaires supplemented with cancer-specific

modules, which combine the general cancer and cancer-specific approaches; and (e) targeted, unidimensional measures, which focus in depth on specific aspects of HRQOL (including symptoms like pain, nausea, and fatigue) and that can be used alone or in combination. Global and preference-based measures (i.e., those that incorporate comparisons of the values attached to different health states) are also reported, although less frequently. Extensive data support the reliability and validity of many of these questionnaires, and recent attention has also focused particularly on the questionnaires’ responsiveness, interpretability, and cultural and language adaptations [see Aaronson (15)]. In contrast, the methodological development of questionnaires to assess patient perceptions of care (including satisfaction) and needs is less mature (10,11).

**4) Outcomes measurement can be applied across the continuum of cancer prevention, detection, survivorship, and care.** As Mandelblatt and Selby (16) point out, individual assessments can be used in cancer prevention and screening, where the emphasis is on not harming healthy individuals (as opposed to treating patients with cancer). Zebrack and Cella (12) note that outcomes assessment is equally applicable in cancer survivorship, because the impact of cancer and its treatment does not cease when treatment ends and these long-term effects of cancer have important consequences on functioning and well-being. Finally, as Ferrell (7) notes, when patients are terminally ill, comfort and well-being should be the primary focus.

**5) Although there is some agreement in the literature about what domains should be included in HRQOL, there is less consensus about which aspects of these domains should be measured.** Ferrans (17) distinguishes between different kinds of HRQOL questions that patients may be asked: questions about objective conditions (e.g., did you go for a long walk today?), questions about perceived status (e.g., could you go for a long walk today?), and evaluation questions (e.g., how satisfied are you with your physical functioning today?). Litwin and Talcott (6) and Darby (10) make similar distinctions regarding questions about symptoms and patient care experiences, respectively. These varying questions—all of which are found in different HRQOL questionnaires—may not necessarily yield consistent findings. For example, patient evaluations may be heavily affected by factors such as expectations, previous experience, and predispositions, whereas these factors may not affect responses to questions about objective conditions. The interrelationships among these different ways of asking questions about HRQOL have not been examined in depth.

**6) The content validity of available HRQOL instrument(s) should be enhanced to optimize their use in particular populations.** For example, Moinpour and Provenzale (4) concluded that social functioning is assessed inadequately in HRQOL measures applied to colorectal cancer patients undergoing treatment. Zebrack and Cella (12) found that current multidimensional HRQOL measures may not capture elements important to cancer survivors, such as fear of recurrence or chronic physical compromise. Williams (18), a prostate cancer survivor, believes that current HRQOL instruments fail to capture the depth of suffering faced by cancer patients and their families.

**7) A number of different HRQOL questionnaires exist, yet few have been compared directly.** This situation raises the proverbial question, “What’s the best HRQOL questionnaire?” Several analytic methods that are currently being developed, particularly item response theory (IRT) modeling, could make this

question moot in the future. IRT modeling is described in more detail below; in brief, IRT-based item banks (that is, compilations of questions) are intended to facilitate the development of computer-adaptive assessments. These assessments would draw from the constellation of available items from many existing HRQOL questionnaires to identify and sequence the items that are most appropriate for each respondent. In theory, IRT applications have the potential to dramatically reduce the need for single (i.e., traditional, fixed-item) instruments.

For the time being, however, COMWG members did not feel that any single HRQOL scale was best for all purposes, because each study requires measures that are appropriate to its distinct hypotheses and patient population. Our summary suggestions for questionnaire selection, which are consistent with the COMWG analyses, are as follows:

- When the goal is to assess HRQOL data at the bedside, the ability of the instrument(s) to provide clear and interpretable data to the clinician is of greatest importance.
- When the research objective is to compare interventions in clinical trials or observational studies, very specific and highly targeted measures may be most appropriate to detect important and clinically significant differences in HRQOL across groups.
- Research questions that investigate patient or provider decision-making may best be addressed through preference-based HRQOL measures, perhaps in conjunction with non-preference-based measures.
- When the research objective is to compare the HRQOL of a study population to that of the population at large (e.g., in studies of cancer prevention or survivorship), a generic HRQOL measure with normative data available on healthy populations is appropriate.
- When the objective is to use data for national policy guidance (e.g., monitoring the cancer burden at the national level to inform budgetary discussions), having a quality-of-life metric that is comparable with those used in other diseases and perhaps in the population more generally may be optimal.

## METHODOLOGIC ACCOMPLISHMENTS AND PROMISING DIRECTIONS

**1) The quantity and quality of HRQOL research in cancer has increased markedly in the past 15 years.** This observation is supported by the emergence of specialty organizations such as the International Society for Quality of Life Research (established in 1990), the creation of journals such as *Quality of Life Research* (established in 1991), and historical analyses [see Spilker (19)]. A MEDLINE search of the period 1990–1994 yielded 2416 articles identified in a search combining the terms “cancer” with “quality of life.” Using the same search terms, 4683 articles were identified for 1995–1999 and 5676 articles were identified for 2000–2003. When using the search term “patient satisfaction” instead of “quality of life,” the numbers of articles identified for each of the three periods were 214, 666, and 970, respectively.

**2) One of the biggest methodologic challenges in cancer HRQOL research is missing data.** If a scheduled HRQOL assessment is missed, the required data cannot be retrieved accurately at some later point or extracted from the medical records. Large amounts of missing data threaten the interpretation of

research studies and in fact have been responsible for closing clinical protocols in the past (20). However, considerable progress has been made in this area. Fairclough (21) discusses a number of techniques to avoid or minimize missing data, including designating responsible staff to manage the HRQOL data collection, training data collectors and study participants, budgeting adequate resources for monitoring and patient follow-up, and considering the adoption of multiple modes of survey administration (e.g., paper-and-pencil self-assessments, in-person interviews, telephone surveys, or electronic [including web-based] approaches). Some of these technologies are just becoming available in the clinical setting and thus their use is only beginning to be reported in the scientific literature.

**3) Approaches to analyzing HRQOL data have undergone considerable development.** Although it is often thought that HRQOL analyses are qualitatively different from biomedical analyses, Sloan (22) argues that the statistical analysis of HRQOL and other patient-reported outcomes poses no greater problem, fundamentally, than is encountered with more traditional biomedical measures. He suggests that it is vital to have a clear statistical modeling plan specified in advance, to pursue complex statistical modeling approaches only after careful basic analyses have been completed, to conduct sensitivity analyses to examine the robustness of study findings as a routine matter, and to present findings clearly and transparently, emphasizing not only summary statistics but graphical displays and other approaches to reveal important variations within the sample.

**4) New approaches to questionnaire development and analysis hold considerable potential for addressing many conceptual and practical concerns.** In particular, the application of IRT to the study of HRQOL is very promising. Although IRT is long established in educational testing research (23–25), it is only recently being applied in the HRQOL arena. IRT provides an alternative perspective to classical test theory (CTT), the more standard approach for assessing and evaluating HRQOL and other patient-reported outcomes that involve latent-variable constructs (i.e., those that can be inferred through indicators but cannot be observed directly). A CTT approach to estimating an individual’s HRQOL level is typically based on the sum of survey item responses, whereas an IRT approach makes use of the pattern of item responses because each item is assumed to convey specific, differentiated information about the individual’s HRQOL level on a particular unidimensional construct.

In essentially tailoring the questionnaire to the individual, IRT modeling offers a number of potential advantages over CTT. First, IRT allows more fine-tuned measurement across the HRQOL continuum, is more efficient, and poses less respondent burden. As noted, it also may lead us to the point that researchers need not choose among competing HRQOL questionnaires, because it is possible that all such items could be pooled, given weights, and even administered as part of a computer-adaptive testing protocol. In such an approach, items would be selected sequentially and strategically based on the respondent’s responses to previous items. IRT modeling also enables a statistically rigorous examination of whether a given instrument performs the same or differently across cultures, geographic borders, and population subgroups through analysis of “differential item functioning.”

Considerable work still needs to be done before IRT can be used in HRQOL assessment. Some concrete issues, such as intellectual property conflicts, may arise with item banking (26). In addition, the multidimensionality of HRQOL can pose challenges

to IRT measurement models, which assume unidimensionality, in instances where a summary measure (single index) is needed. It should also be noted that CTT imposes a similar unidimensionality assumption. Also, IRT's application may be limited unless the technology to support these methods (e.g., computer-adaptive technology) is available. Nevertheless, there was virtual consensus among COMWG participants that IRT provides the most (if not the only) sound theoretical basis for item banking, computer-adaptive testing, cross-walking scale scores (that is, comparing items from different questionnaires) via item linking, and investigating differential item functioning. This area is developing fast [for an update, see reference (27)].

**5) Additional analytic approaches, such as structural equation modeling and Bayesian analysis, may be useful in HRQOL research.** Structural equation modeling may offer a useful platform for exploring issues related to conceptual models of HRQOL. This approach facilitates the development of causal models that depict relationships among sets of survey items, scales and subscales, other patient-reported (e.g., symptoms) and biomedical (e.g., toxicity) outcomes, and exogenous factors, such as respondent age, race, sex, and educational level. To date, there have been comparatively few applications of structural equation modeling to HRQOL (28); future analysis will clarify whether this approach will help to identify causal indicators for HRQOL.

As O'Brien (29) argues, Bayesian statistical modeling has been underused in HRQOL research to date. Such techniques may be particularly useful in cost-effectiveness and cost-utility analyses that take a lifetime perspective.

## USE OF HRQOL ASSESSMENTS AND INTERPRETATION OF FINDINGS

**1) The primary areas in which outcomes assessments, and most often HRQOL measurement, have been used to date are clinical trials and observational studies of cancer treatment.** The most frequent research questions have concerned determining the intervention that is preferable in terms of HRQOL in randomized clinical trials and identifying short and long-term effects of cancer therapy. Fewer studies have reported HRQOL assessments in clinical care, surveillance, or policy decision making. For example, Osoba (30) discusses how HRQOL assessment can be used to facilitate patient-physician communication. The Centers for Medicare & Medicaid Services has recently initiated a demonstration project that will reimburse physicians for using standardized assessments of pain, nausea and vomiting, and fatigue among cancer patients undergoing chemotherapy (31). Donaldson has also identified ways that new technologies and changes in the health care system could facilitate the incorporation of HRQOL and other patient-reported outcomes into routine clinical practice (32).

**2) There is much more research available about HRQOL in some cancers than in others.** This situation affects the conclusions that can be drawn.

In breast cancer, many studies that include HRQOL assessment have been published, partially due to the relatively high prevalence of this diagnosis and breast cancer survivors. Ganz and Goodwin (3) concluded that younger patients and those with more advanced disease report more distress, that survivors generally have HRQOL equivalent to that in control groups, that HRQOL improves in the year postdiagnosis and worsens with

recurrence, and that psychological interventions confer substantial benefits. These authors also found, surprisingly, that HRQOL in breast cancer patients receiving adjuvant chemotherapy has been studied only infrequently and that HRQOL data have not contributed to the interpretation of most of the biomedical breast cancer treatment studies reported to date.

The analyses for each of the other three cancer sites (lung, colorectal, and prostate cancer) all yielded somewhat different conclusions from the breast cancer analyses and from each other. There was considerably less literature available for review in lung and colorectal cancer than in breast and prostate cancer. The existence of relatively few HRQOL studies in these cancers may partially reflect the fact that only 39% of colorectal cancers and 16% of lung cancers are diagnosed when they are still localized (33). HRQOL assessment in more advanced disease presents distinct methodologic challenges from its assessment in localized disease, particularly with respect to missing data in prospective studies. In both lung and colorectal cancer, studies of both primary and adjuvant treatments and their association with HRQOL are common. Earle and Weeks (5) noted that, in lung cancer, health care professionals believe that patients' HRQOL is worse than the patients think it is. In prostate cancer, a disease for which the current therapeutic modalities for early-stage disease have long-term consequences for sexual and bladder function, measures of impotence, incontinence, and other symptoms were frequently reported (6). Litwin and Talcott (6) noted that HRQOL data have shown that patient impairments in these areas are much more common than clinicians have realized.

**3) The interpretation of HRQOL findings is limited by the absence of widely accepted models to define HRQOL and to specify predictive and correlated variables.** As Ferrans (17) discusses, HRQOL has yet to be consistently defined; consequently, different and sometimes competing definitions may be used. Darby (10) points out that the same is true for patient satisfaction. Erickson (9) observes that even when two measures purport to assess the same domains, they may use very different items to measure what is theoretically the same concept. Because HRQOL questionnaires vary widely in their domains and items, much confusion remains as to what HRQOL really is.

The conceptual underpinnings of measurement models also need additional attention. For HRQOL (and also patient perceptions of care and needs assessment), conceptual models are required that include a theory that specifies which domains are to be included in the HRQOL measurement model, the relationships among the various domains, the relative importance of the domains, and a causal pathway that clearly distinguishes causal from indicator variables [in the terminology of Fayers and Machin (28)] and that specifies the relationships between them.

One important factor that needs to be considered in such models is "response shift." As discussed by Schwartz and Sprangers (34,35), this phenomenon refers to how a patient's changing internal standards and values affect HRQOL ratings over time. As Ferrans (17) emphasizes, a deeper understanding of whether response shift influences HRQOL assessments in oncology could open the way to conceptual models that better account for the complex relationship between changes in relatively objective biomedical outcomes and more malleable HRQOL measures.

**4) There is an emerging consensus for how to interpret HRQOL scores.** Osoba (30), Sloan (22), and a recent series of meetings sponsored by the Mayo Clinic (36) have provided insights about clinical interpretability. Considerable progress has

been made in terms of quantifying what patients perceive to be meaningful changes in HRQOL, as reflected in their scores on HRQOL questionnaires, and the extent to which such scores are associated with differences that make sense to clinicians. Osoba's evidence-based conclusion that a small, perceptible, and meaningful change in an HRQOL score appears to be about 7% of the full scale breadth (perhaps bracketed by 5% and 10%) is a genuinely surprising if not felicitous finding that bears ongoing investigation by anchor-based approaches, as he rightly emphasizes (30). Among such anchors should be those that have what we might term "decision significance," i.e., a relationship between observed changes in HRQOL scores and choices, real or hypothetical, made by patients or other decision makers (37).

**5) The "value added" by HRQOL assessments, particularly in clinical trials with a survival endpoint, is not well understood.** In a number of trials, HRQOL is the primary endpoint of the study; as such, HRQOL data provide the major indications of intervention efficacy in these studies. However, in trials whose primary endpoint is survival and for which HRQOL is a secondary endpoint, the contributions of HRQOL data are not always clear. If HRQOL and survival covary (e.g., if survival and HRQOL both increase), the HRQOL data may be regarded by some trialists as redundant. Given the significance of increasing survival for both clinicians and patients (38,39), HRQOL data may have a limited impact on study interpretation when survival differences are found. In the words of some COMWG members, "Survival trumps all."

In instances when HRQOL data do not covary with survival data, several explanations are possible (4,40). The HRQOL measurement tool may not have been sensitive enough to detect real differences in patient well-being; the HRQOL measurement tool may not have focused on the right aspects of HRQOL; the timing of the HRQOL assessment may not have been appropriate to when patients perceived changes in HRQOL; the trial may have had methodologic flaws that affected the HRQOL data, such as small sample size or large amounts of missing data; or the (statistically significant) difference in biomedical outcomes did not in fact correspond to a clinical benefit (or harm) that was both perceptible and important to the patient, which would be consistent with the observation that there was no clinically meaningful change in HRQOL.

Given the multiple possible explanations for any set of findings, we believe that it is premature to conclude, based on clinical trials, that HRQOL data do or do not contribute to treatment decision making. Careful attention needs to be given to the selection of HRQOL measures for clinical trials and to understanding the relationships between and among the various outcome measures. In particular, the relationship between symptom assessment and HRQOL should be explored. This analysis is especially important considering that the Food and Drug Administration has based its regular approval of 13 out of 57 new cancer agents on the relief of tumor-specific symptoms (41).

## NEXT STEPS

This overview—based on a bird's-eye view of the in-depth work carried out by COMWG members—shows that cancer outcomes measurement has come a long way in a relatively short period of time. Measures are available, they have been used and refined, and further methodologic advances on the horizon promise a qualitative increase in the sophistication of measurement

techniques in the future. However, for cancer outcomes research to contribute optimally to the goal of reducing death and suffering due to cancer, additional progress is needed. In particular, we need sound evidence about the impact of interventions on such outcomes and we must develop both the capacity and the commitment to translate this evidence (about efficacy, effectiveness, and cost-effectiveness) to findings that will be useful to the full spectrum of decision makers.

Several kinds of future activities may facilitate this translation. These might include:

- A state-of-the-science meeting to take stock of COMWG findings and communicate them to a broader audience, to examine patient-reported outcomes in cancer disease sites beyond the four examined here, to lay the groundwork for formal consensus development, and to formulate a research agenda for the outcomes measurement field.
- Continued building of research capacity to address the many issues in this area, only some of which have been discussed in this commentary. The National Institutes of Health has already made commitments to this field, including a recent \$25 million, 5-year extramural research project to support innovative applications of IRT, including item banks and computer-adaptive testing for chronic diseases like cancer (42). Other creative mechanisms might also be considered, such as one suggested by COMWG members: the creation of cancer outcomes research teams to conduct multi- and interdisciplinary studies on a range of topics while also training the next generation of cancer outcomes researchers.
- Enhancing the accuracy, timeliness, and "linkability" of the major data sources: cancer registries; medical records; administrative files including claims information; and surveys of patients, individuals at risk of cancer, and health care providers (43). Advances in information technology are expected to accelerate the adoption of electronic health records. In addition, NCI's efforts, such as caBIG (Cancer Biomedical Informatics Grid), may help facilitate such linkages in the future. Although issues related to the protection of patient privacy must continue to be addressed, there has never been a more opportune time to encourage the creation, testing, and expansion of linked data systems to support cancer outcomes monitoring, evaluation, and improvement.
- Investigating the actual and potential roles of outcome measures in cancer care decision making, focusing on the patient and family, providers, payers, regulators, and other policy makers, including the Food and Drug Administration. A multifaceted research strategy of decision-making processes may be required that includes feasibility studies, field surveys, analyses of previous decisions including case studies, focus groups, and in-depth interviews of decision participants.
- Although the COMWG's major emphasis was on HRQOL outcomes, the above recommendations apply equally to the assessment of patient satisfaction and economic outcomes. Strengthening the knowledge base in these areas is essential (10,11,29,44).

Reducing deaths and suffering due to cancer is a compelling and daunting challenge that is being pursued by a worldwide community of cancer researchers, providers, patients, survivors, families, caregivers, and volunteers. With basic and clinical scientific discoveries now accelerating the development of promising new

interventions and with a heightened public and private focus on delivering quality cancer care to all who need it, there are good reasons to expect substantial and sustained progress in reducing the cancer burden. Through cancer outcomes assessment, we can monitor the progress being achieved at any point in time while conducting the research needed to inform decisions that will have a direct impact on reducing the burden over time.

#### APPENDIX: MEMBERS OF THE NCI CANCER OUTCOMES MEASUREMENT WORKING GROUP, 2001–2004

**Cochairs:** Joseph Lipscomb, PhD (Chief, Outcomes Research Branch, Applied Research Program, Division of Cancer Control and Population Sciences, National Cancer Institute); Carolyn C. Gotay, PhD (Professor, Cancer Research Center of Hawaii, University of Hawaii); **Working Group Initiator:** Claire Snyder, MHS (Expert, Outcomes Research Branch, Applied Research Program, Division of Cancer Control and Population Sciences, National Cancer Institute); **Working Group Participants:** Neil K. Aaronson, PhD (Head, Division of Psychosocial Research & Epidemiology, The Netherlands Cancer Institute and Professor, Faculty of Medicine, Vrije Universiteit); Michael J. Barry, MD (Chief, General Medicine Unit, Massachusetts General Hospital); David Cella, PhD (Professor of Psychiatry and Behavioral Science, Northwestern University Feinberg School of Medicine and Director, Center on Outcomes Research and Education, Evanston Northwestern Healthcare); Janet E. Dancey, MD (Senior Clinical Investigator, Investigation Drug Branch, Cancer Therapy Evaluation Program, Division of Cancer Treatment and Diagnosis, National Cancer Institute); Charles Darby (Social Science Administrator, Agency for Healthcare Research and Quality); Craig C. Earle, MD, MSc (Assistant Professor of Medicine, Harvard Medical School, Dana-Farber Cancer Institute); Pennifer Erickson, PhD (Associate Professor, Departments of Biobehavioral Health and Health Evaluation Sciences, Pennsylvania State University); Diane L. Fairclough, DrPH (Professor, Colorado Health Outcomes Center and Department of Preventive Medicine and Biometry, University of Colorado Health Sciences Center); David H. Feeny, PhD (Professor of Pharmacy and Pharmaceutical Sciences, Departments of Economics and Public Health Sciences, University of Alberta); Carol Estwing Ferrans, PhD, RN, FAAN (Professor, College of Nursing, University of Illinois at Chicago); Betty R. Ferrell, PhD, FAAN (Research Scientist, City of Hope Medical Center); Patricia A. Ganz, MD (Professor, Schools of Medicine and Public Health, and Director, Division of Cancer Prevention and Control Research, Jonsson Comprehensive Cancer Center, University of California, Los Angeles); Pamela J. Goodwin, MD, MSc, FRCP (C) (Senior Scientist, Samuel Lunenfeld Research Institute, Mount Sinai Hospital, Professor of Medicine, University of Toronto); David H. Gustafson, PhD (Robert Ratner Professor of Industrial Engineering and Director, Center of Excellence in Cancer Communications Research, University of Wisconsin, Madison); Ronald K. Hambleton, PhD (Distinguished University Professor, School of Education, University of Massachusetts); Mark C. Hornbrook, PhD (Chief Scientist, Center for Health Research, Northwest and Hawaii, Kaiser Permanente, Northwest Region); Mark S. Litwin, MD, MPH (Professor of Urology and Health Services, Schools of Medicine and Public Health, University of California, Los Angeles); Jeanne S. Mandelblatt, MD, MPH (Director, Cancer &

Aging Research, Lombardi Comprehensive Cancer Center and Departments of Oncology and Medicine, Georgetown University Medical Center); Mary S. McCabe, RN, MA (Director, Office of Education and Special Initiatives, National Cancer Institute); Carol M. Moinpour, PhD (Behavioral Scientist, Southwest Oncology Group Statistical Center and Associate Member, Division of Public Health Sciences, Fred Hutchinson Cancer Research Center); Bernie J. O'Brien, PhD (Professor, Department of Clinical Epidemiology and Biostatistics, McMaster University and Associate Director, Centre for Evaluation of Medicines, St. Joseph's Healthcare); David Osoba, BSc, MD, FRCPC (Quality of Life Consultant, QOL Consulting, West Vancouver, BC); Dawn Provenzale, MD, MS (Associate Professor of Medicine and Director GI Outcomes Research, Duke University Medical Center); Steven P. Reise, PhD (Professor, Department of Psychology, University of California, Los Angeles); Dennis A. Revicki, PhD (Vice President and Director, Center for Health Outcomes Research, MEDTAP International); Joe V. Selby, MD, MPH (Director, Division of Research, Kaiser Permanente Northern California); Jeff A. Sloan, PhD (Lead Statistician, Cancer Center Statistics, Mayo Clinic Rochester); James A. Talcott, MD, SM (Assistant Professor and Director, Center for Medical Outcomes, Massachusetts General Hospital); Jane C. Weeks, MD, MSc (Associate Professor of Medicine and Chief, Division of Population Science, Dana-Farber Cancer Institute); James E. Williams, Jr. (Col. Ret., USA) (Co-Chairman, Pennsylvania Prostate Cancer Coalition and Vice President, Intercultural Cancer Council Caucus); Mark Wilson, PhD (Professor, Graduate School of Education, University of California, Berkeley); and Brad Zebrack, PhD, MSW (Cancer Survivor/Advocate, NCI Director's Consumer Liaison Group and Assistant Professor, School of Social Work, University of Southern California).

#### REFERENCES

- (1) Lipscomb J, Gotay CC, Snyder C, editors. Outcomes assessment in cancer. Cambridge (United Kingdom): Cambridge University Press; 2005.
- (2) Lohr KN. Assessing health status and quality-of-life instruments: attributes and review criteria. *Qual Life Res* 2002;11:193–205.
- (3) Ganz PA, Goodwin PJ. Quality of life in breast cancer—what have we learned and where do we go from here? In: Lipscomb J, Gotay CC, Snyder C, editors. Outcomes assessment in cancer. Cambridge (United Kingdom): Cambridge University Press; 2005. p. 93–125.
- (4) Moinpour CM, Provenzale D. Treatment for colorectal cancer: impact on health-related quality of life. In: Lipscomb J, Gotay CC, Snyder C, editors. Outcomes assessment in cancer. Cambridge (United Kingdom): Cambridge University Press; 2005. p. 178–200.
- (5) Earle CC, Weeks JC. The science of quality-of-life measurement in lung cancer. In: Lipscomb J, Gotay CC, Snyder C, editors. Outcomes assessment in cancer. Cambridge (United Kingdom): Cambridge University Press; 2005. p. 160–77.
- (6) Litwin MS, Talcott JA. Measuring quality of life in prostate cancer: progress and challenges. In: Lipscomb J, Gotay CC, Snyder C, editors. Outcomes assessment in cancer. Cambridge (United Kingdom): Cambridge University Press; 2005. p. 126–59.
- (7) Ferrell BR. Assessing health-related quality of life at end of life. In: Lipscomb J, Gotay CC, Snyder C, editors. Outcomes assessment in cancer. Cambridge (United Kingdom): Cambridge University Press; 2005. p. 264–85.
- (8) Snyder C. Assessing the subjective impact of caregiving on informal caregivers of cancer patients. In: Lipscomb J, Gotay CC, Snyder C, editors. Outcomes assessment in cancer. Cambridge (United Kingdom): Cambridge University Press; 2005. p. 329–45.
- (9) Erickson P. Assessing health status and quality of life of cancer patients: the use of general instruments. In: Lipscomb J, Gotay CC, Snyder C, editors.

- Outcomes assessment in cancer. Cambridge (United Kingdom): Cambridge University Press; 2005. p. 31–68.
- (10) Darby C. Measuring the patient's perspective on the interpersonal aspects of cancer care. In: Lipscomb J, Gotay CC, Snyder C, editors. Outcomes assessment in cancer. Cambridge (United Kingdom): Cambridge University Press; 2005. p. 290–304.
- (11) Gustafson DH. Needs assessment in cancer. In: Lipscomb J, Gotay CC, Snyder C, editors. Outcomes assessment in cancer. Cambridge (United Kingdom): Cambridge University Press; 2005. p. 305–28.
- (12) Zebrack B, Cella D. Evaluating quality of life in cancer survivors. In: Lipscomb J, Gotay CC, Snyder C, editors. Outcomes assessment in cancer. Cambridge (United Kingdom): Cambridge University Press; 2005. p. 241–63.
- (13) Potosky AL, Harlan LC, Stanford JL, Gilliland FD, Hamilton AS, Albertsen PC, et al. Prostate cancer practice patterns and quality of life: the Prostate Cancer Outcomes Study. *J Natl Cancer Inst* 1999;91:1719–24.
- (14) Gotay CC. Cancer-related quality of life across the spectrum of applications. *J Natl Cancer Inst Monogr* 2004;33:126–33.
- (15) Aaronson NK. Cross-cultural use of health-related quality of life assessments in clinical oncology. In: Lipscomb J, Gotay CC, Snyder C, editors. Outcomes assessment in cancer. Cambridge (United Kingdom): Cambridge University Press; 2005. p. 406–24.
- (16) Mandelblatt JS, Selby JV. Short-term outcomes of chemoprevention, genetic susceptibility testing, and screening interventions: What are they? How are they measured? When should they be measured? In: Lipscomb J, Gotay CC, Snyder C, editors. Outcomes assessment in cancer. Cambridge (United Kingdom): Cambridge University Press; 2005. p. 216–40.
- (17) Ferrans CE. Definitions and conceptual models of quality of life. In: Lipscomb J, Gotay CC, Snyder C, editors. Outcomes assessment in cancer. Cambridge (United Kingdom): Cambridge University Press; 2005. p. 14–30.
- (18) Williams JE. Patient advocate perspective on health-related quality of life issues with prostate cancer survivors. In: Lipscomb J, Gotay CC, Snyder C, editors. Outcomes assessment in cancer. Cambridge (United Kingdom): Cambridge University Press; 2005. p. 286–89.
- (19) Spilker B. The world of outcomes research: yesterday, today, and tomorrow. In: Lipscomb J, Gotay CC, Snyder C, editors. Outcomes assessment in cancer. Cambridge (United Kingdom): Cambridge University Press; 2005. p. 584–89.
- (20) Hayden KA, Moinpour CM, Metch B, Feigl P, O'Bryan RM, Green S, et al. Pitfalls in quality-of-life assessment: lessons from a Southwest Oncology Group breast cancer clinical trial. *Oncol Nurs Forum* 1993;20:1415–9.
- (21) Fairclough DL. Practical considerations in outcomes assessment for clinical trials. In: Lipscomb J, Gotay CC, Snyder C, editors. Outcomes assessment in cancer. Cambridge (United Kingdom): Cambridge University Press; 2005. p. 346–61.
- (22) Sloan JA. Statistical issues in the application of cancer outcomes measures. In: Lipscomb J, Gotay CC, Snyder C, editors. Outcomes assessment in cancer. Cambridge (United Kingdom): Cambridge University Press; 2005. p. 362–85.
- (23) Bock, RD. A brief history of item response theory. *Educational measurement: issues and practice* 1997;16:21–33.
- (24) Lord, FM. Applications of item response theory to practical testing problems. Hillsdale (NJ): Lawrence Erlbaum Associates; 1980.
- (25) Wright, BD. A history of social science measurement. *Educational measurement: issues and practice* 1997;16:33–45.
- (26) McHorney CA, Cook K. The ten D's of health outcomes measurement for the 21st century. In: Lipscomb J, Gotay CC, Snyder C, editors. Outcomes assessment in cancer. Cambridge (United Kingdom): Cambridge University Press; 2005. p. 590–609.
- (27) Item Banking Special Issue. *Quality of Life* newsletter. Lyon (France): MAPI Research Trust; Fall 2004.
- (28) Fayers PM, Machin D. *Quality of life: assessment, analysis and interpretation*. Chichester (Great Britain): John Wiley & Sons Ltd.; 2002.
- (29) O'Brien BJ. Cost-effectiveness analysis in cancer: toward an iterative framework for integration of evidence from trials and models. In: Lipscomb J, Gotay CC, Snyder C, editors. Outcomes assessment in cancer. Cambridge (United Kingdom): Cambridge University Press; 2005. p. 503–21.
- (30) Osoba D. The clinical value and meaning of health-related quality-of-life outcomes in oncology. In: Lipscomb J, Gotay CC, Snyder C, editors. Outcomes assessment in cancer. Cambridge (United Kingdom): Cambridge University Press; 2005. p. 386–405.
- (31) Centers for Medicare & Medicaid Services. Demonstration of improved quality of care for cancer patients undergoing chemotherapy. Available at: <http://www.cms.hhs.gov/media/press/release.asp?Counter=1245>. [Last accessed: October 3, 2005.]
- (32) Donaldson, MS. Taking stock of health-related quality-of-life measurement in oncology practice in the United States. *J Natl Cancer Inst Monogr* 2004;33:155–67.
- (33) American Cancer Society. *Cancer facts & figures, 2005*. Atlanta (GA): American Cancer Society; 2005.
- (34) Schwartz CE, Sprangers MAG, editors. *Adaptation to changing health: response shift in quality-of-life research*. Washington, DC: American Psychological Association; 2000.
- (35) Sprangers MAG, Schwartz CE. Integrating response shift into health-related quality of life research: a theoretical model. *Soc Sci Med* 1999;48:1507–15.
- (36) Sloan JA, Cella D, Frost MH, Guyatt GH, Sprangers M, Symonds T, et al. Assessing clinical significance in measuring oncology patient quality of life: introduction to the symposium, content overview, and definition of terms. *Mayo Clin Proc* 2002;77:367–70.
- (37) Dowie J. Decision validity should determine whether a generic or condition-specific HRQOL measure is used in health care decisions. *Health Econ* 2002;11:1–8.
- (38) Lindley C, Vasa S, Sawyer WT, Winer EP. Quality of life and preferences for treatment following systemic adjuvant therapy for early-stage breast cancer. *J Clin Oncol*. 1998;16:1380–7.
- (39) Slevin ML, Stubbs L, Plant HJ, Wilson P, Gregory WM, Armes PJ, et al. Attitudes to chemotherapy: comparing view of patients with cancer with those of doctors, nurses, and general public. *BMJ* 1990;300:1458–60.
- (40) Barry MJ, Dancy JE. Instruments to measure the specific health impact of surgery, radiation, and chemotherapy on cancer patients. In: Lipscomb J, Gotay CC, Snyder C, editors. Outcomes assessment in cancer. Cambridge (United Kingdom): Cambridge University Press; 2005. p. 201–15.
- (41) Johnson JR, Williams G, Pazdur R. End points and United States Food and Drug Administration approval of oncology drugs. *J Clin Oncol* 2003;21:1404–11.
- (42) National Institutes of Health. NIH Roadmap Initiative RFA-AR-04-007. Available at: <http://grants1.nih.gov/grants/guide/rfa-files/RFA-RM-04-011.html>. [Last accessed: December 25, 2003.]
- (43) Gotay CC, Lipscomb J. Data for cancer outcomes research: identifying and strengthening the empirical base. In: Lipscomb J, Gotay CC, Snyder C, editors. Outcomes assessment in cancer. Cambridge (United Kingdom): Cambridge University Press; 2005. p. 522–49.
- (44) Hornbrook, MC. On the definition and measurement of the economic burden of cancer. In: Lipscomb J, Gotay CC, Snyder C, editors. Outcomes assessment in cancer. Cambridge (United Kingdom): Cambridge University Press; 2005. p. 480–502.

## NOTES

Joseph Lipscomb's former affiliation was Chief, Outcomes Research Branch, Applied Research Program, Division of Cancer Control and Population Sciences, National Cancer Institute, Bethesda, MD. Claire F. Snyder's former affiliation was Expert, Outcomes Research Branch, Applied Research Program, Division of Cancer Control and Population Sciences, National Cancer Institute, Bethesda, MD.

The views expressed are those of the authors and do not necessarily represent the position of the National Cancer Institute.

Manuscript received April 20, 2005; revised August 26, 2005; accepted September 8, 2005.