

Preliminary Report: The Development of the NCCN Comparative Therapeutic Index™ as a Clinical Evaluative Process for Existing Data in Oncology

Edward C. Li, PharmD, BCOP, and Jessica DeMartino, PhD,* Fort Washington, Pennsylvania

Key Words

NCCN, comparative therapeutic index, comparative effectiveness research, evidence-based medicine, guidelines

Abstract

The National Comprehensive Cancer Network (NCCN) develops and communicates the NCCN Clinical Practice Guidelines in Oncology (NCCN Guidelines) to oncologists and other clinicians. The NCCN Guidelines are widely recognized and applied as the standard for clinical policy in the United States. These guidelines and related documents, such as the NCCN Drugs & Biologics Compendium (NCCN Compendium), are used extensively by public and private payors as the basis for the setting of coverage policies. Given the demand for comparative effectiveness (CE) analyses, as described and discussed in this report, the NCCN has begun work on a paradigm to integrate evidence-based CE analysis into the NCCN Guidelines deliberative process. This report presents NCCN's initial thinking on the use of NCCN expert panel members in developing a process that can be used to compare health care technologies (e.g., radiation modalities, chemotherapy regimens) in a formal, systematic way. Draft considerations are provided to stimulate discussion and feedback, particularly in the oncology community, as NCCN moves through processes such as methodologic review, validation of rating scales, and review of implications for public policy, toward finalization of an NCCN CE analytic paradigm. (*JNCCN* 2010;8[Suppl 5]:S1–S9)

The increasing availability of innovative health care technologies for the treatment of cancer and the substantial expense of some of these technologies are driving the need for a more explicit comparison of therapeutic options for specific clinical indications. Major constituencies, including legislators and policy makers, are calling for the application of comparative effectiveness (CE) analysis in developing clinical policy. CE research (CER) has emerged as a priority concept as part of the larger health care reform agenda. The need for the attainment of optimal value for the dollars spent on health care has become acute. National health expenditures are approaching \$2.5 trillion and 17% of the gross domestic product.^{1,2} The pressure on the United States health care system will only become greater as health care reform will likely add substantial numbers of uninsured individuals to the rolls of those insured for the receipt of health services.³

Increasingly, there has been recognition of “knowledge gaps” regarding the most effective intervention or treatment for a given indication. Because of the current regulatory climate and approval process, newly introduced technologies are not always compared head-to-head against the other available options, particularly the leading available intervention. The resultant uncertainty about the most effective and appropriate therapy most likely factors into the occurrence of inconsistent care and variable outcomes within our current health care system.⁴ CER seeks to address these issues by obtaining additional information regarding the relative effectiveness of different treatment options. CER should be the genesis of new information, or synthesis and analysis of

From the National Comprehensive Cancer Network, Fort Washington, Pennsylvania.

*This report was developed in conjunction with a working group of NCCN physicians and outside experts.

The authors have disclosed that they have no financial interests, arrangements, or affiliations with the manufacturers of any products discussed in this article or their competitors; they are employees of the National Comprehensive Cancer Network.

Correspondence: Edward C. Li, PharmD, BCOP, National Comprehensive Cancer Network, 275 Commerce Drive, Suite 300, Fort Washington, PA 19034.

E-mail: li@nccn.org

existing data, that can comparatively evaluate all facets of health care, from diagnosis to treatment, for the purpose of allowing relevant stakeholders (e.g., providers and patients) to make the best decision about what is effective and efficient care. Although the definition of CER has already been established, the strategies surrounding the proper conduct of CER projects and the implementation of these results into routine clinical practice are the current focus of discussion.

The need for better use of existing research methods, development of innovative research paradigms, and more formal and effective use of existing data has been recognized by many national committees and organizations. In early 2009, the federal government allocated \$1.1 billion in funding in the American Recovery and Reinvestment Act (ARRA) for CER.⁵ These dollars are directed to government agencies such as the Agency for Healthcare Research and Quality (AHRQ), the National Institutes of Health (NIH), and the Department of Health and Human Services (HHS) for programmatic administration. As part of this legislation, a Federal Coordinating Council for Comparative Effectiveness Research (FCCCER) was established to “coordinate and guide research investments in comparative effectiveness research funded by the Recovery Act.” Additionally, the health care reform bill, “Patient Protection and Affordable Care Act,” increased provisions for CER through increasing funding and establishing a nonprofit entity known as the Patient-Centered Outcomes Research Institute (PCORI) to identify research priorities and conduct research that compares the clinical effectiveness of medical treatments.⁶ Once PCORI is in place, the FCCCER will be terminated.

The enacted legislation underscores the importance of CER in current and future U.S. health policy. In discussions about the rapid pace of scientific advancement coupled with expensive health care technologies, no area of medicine is afforded more emphasis than oncology. As payors and other constituencies of the health care community focus on cancer care, it is critical that the oncology community be actively involved in the development and implementation of processes at all levels.

To address the knowledge gap problem with CER, consideration must be given to the current challenges of transforming CER ideas into improved

patient outcomes, and develop a strategy for overcoming these challenges. The Friends of Cancer Research White Paper (FWP) emphasized the preeminence of randomized controlled clinical trials as the standard for the development of scientific data that defines the safety and effectiveness of a health care technology.⁷ Additionally, the FWP affirmed the value of large outcomes databases or registries in addressing important clinical issues and providing scientific data for comparative analyses.

Although randomized clinical trials (RCTs) are recognized as the gold standard for determining safety and efficacy, their universal application in making CER determinations has challenges. First, RCTs usually focus on a narrowly selected population in a controlled setting and may use surrogate end points.⁸ Selectively enrolling patients allows for robust internal validity, but the results cannot always be readily generalized and extrapolated to the entire patient population likely to receive the treatment. Additionally, RCTs may not be the most practical approach to CER, because they are expensive and require significant time investment. One proposed alternative to RCTs is the use of “practical clinical trials,” wherein the patient population being studied reflects those most likely to receive the treatment in routine clinical practice, and the outcomes studied are those most relevant to clinical decision-making.⁹ Other alternatives to RCTs include nonrandomized observational and/or retrospective analyses of registries, claims data, or other types of databases (i.e., secondary data sources).¹⁰ Importantly, nonrandomized studies of secondary data sources have different strengths and limitations from RCTs, and concern exists regarding their methodologic conduct and interpretation.

Reports such as the one from the International Society for Pharmacoeconomics and Outcomes Research (ISPOR) emphasize the value of data from retrospective database analysis based on good research practices. The 3-part ISPOR report provides guidance on how to design a CER study from secondary data sources that minimizes bias and confounding variables, and uses analytic techniques to infer causality.¹¹⁻¹³ Furthermore, guidance is provided on how to interpret the results derived from these types of studies. CER using secondary data sources would require a significant investment in the current data collection infrastructure to ensure quality of data. Secondary data sources would likely be derived from

registries, claims databases, or electronic health records currently used in the day-to-day practice of caring for patients.

In addition to these challenges regarding methodology and different data sources, with emphasis on the use of existing data, the literature has further expanded on some of the challenges to establishing formal CER programs in oncology. Particular concerns and challenges may include:

- Standards for taxonomy and methodology for CER have not yet been clearly defined or established within the context of oncology.^{2,10,14}
- Given the vast number of “information gaps” in current medical literature pertaining to cancer, prioritization of research questions specific to cancer is necessary.^{2,15}
- CE programs should be continuously evaluated to measure their impact on policies and practices and to identify means of improving dissemination of information.^{2,16}
- CER in oncology is not necessarily limited to comparisons of active cancer treatment. The implementation of different health strategies could also be compared.¹⁶
- “Individualized (or personalized) medicine” is a growing field in oncology and could be advanced through CER by analyzing different subgroups of patients receiving a specific treatment.^{17,18}
- Translation, adoption, and dissemination of results from available CE studies have been limited. An ideal strategy for disseminating this research in a manner that can influence clinical decision-making must combine valid evidence and appropriate interpretation of data.^{2,19}
- A systematic framework for incorporating the existing body of evidence to comparatively evaluate treatment strategies should be developed to facilitate the selection of appropriate interventions while accounting for patient-specific comorbidities and situations.^{2,19}

The nation’s great biomedical research enterprise that is based on the RCT will continue to generate data critical to advance all areas of medicine. However, a need exists to better address decision-making processes today with improved use of the data, information, and expert judgment that are currently available.

Incorporating CE Results Into Practice: Using a Clinical Evaluative Process

Most will concur with the stated goal of CER, which is to assist health care providers, patients, and other stakeholders make informed decisions regarding the provision of care.²⁰ None can deny that increasing the amount of data and information on various treatment options will benefit patients in the long-term. Although the current debate focuses on developing standards for conducting this research, what is missing from the discussion is how these CER results will be translated into clinical practice recommendations.

The proper translation of CER results into clinical practice is a necessary last step in the evidence-based medicine process of caring for patients. The clinical application of these studies could considerably aid the clinical decision-making process, which currently judges and evaluates the available data through indirect comparisons. However, even well-designed CE studies will still have flaws, such as the concerns about whether the study is timely given the rapid pace of changing practice standards, the clinical end points assessed were relevant, and the results are clinically significant for specific subpopulations. Recognizing that these flaws exist makes a final clinical appraisal by experts, based on the whole body of literature and integrating clinical experience, even more important.

Furthermore, ideal CER studies are unlikely to be available in the foreseeable future. The rapidity of knowledge advancement, especially in oncology, and deficiencies in data infrastructure and availability render impractical the sole reliance on these studies for practice guidance. The clinical decision-making realm, as NCCN is attuned to, acknowledges the limitations of using available studies and their corresponding data for direct comparison. However, if patient care is to advance, a judgment must be made using this evidence to synthesize an appraisal as best directed by these clinical comparisons.

It is important to note that a clinical evaluation of the literature along with an informal comparison of the treatment options already occurs whenever treatment decisions are made. Individual clinicians, in their own minds, have made judgments on the therapeutic indices (i.e., effectiveness vs. toxicity) of available options based on available data and personal clinical experience, and have subsequently

compiled a comparative list to be applied to their patients. The final selection of a specific treatment is then made based on patient-specific parameters. NCCN proposes a draft paradigm, named the *Comparative Therapeutic Index (CTI)*, for consideration to systematize the above judgment in an explicitly stated manner, using the expert judgment of oncology thought leaders from NCI-designated cancer centers. As with all NCCN recommendations, these judgments would not be prescriptive, and the final selection of the specific treatment is the responsibility of the individual physician based on patient-specific parameters elucidated during the course of the physician-patient relationship.

Developing the Scoring Tools for the CTI

The NCCN CTI as a clinical evaluative process would integrate the complete body of existing data, clinical experience, and expert judgment when comparing different treatment options. When considering how to establish this process, a few concepts emerged. First, the end product (i.e., the CTI) should be a simple, user-friendly depiction of the aggregate judgments by expert panel members regarding the relative effectiveness and toxicity of different treatment options. Second, the parameters according to which these judgments are made should be explicit and determined *a priori*. Third, the scoring tools used to capture these judgments must take into consideration the clinical decision-making process and reflect how physicians and other clinicians make their determinations for practice. Fourth, the process of developing and validating these parameters and scoring tools should be transparent and collaborative with organizations external to NCCN early in the process. Lastly, the process must be sufficiently facile and efficient that it can be readily integrated into the current NCCN Clinical Practice Guidelines in Oncology (NCCN Guidelines) development process.

The development of the scoring tools for the CTI would be modeled after the methods stated in the FDA's guidance for developing patient-reported outcomes.²¹ In the guidance statement, the FDA describes a process for developing and modifying scoring instruments, including assessment of reliability and validity. Although the guidance document focuses on tools for patients to complete, the methods for determining principal concepts, creating the scoring tools, assessing the reliability and validity, and modifying the instrument are relevant to the CTI process. Therefore, the development of the NCCN CTI encompasses the creation of expert working groups to first develop the conceptual framework and scoring tools, pilot these instruments, and subsequently assess the reliability and validity of the scoring tools and modify them as needed.

Conceptually, a rating scale that judges the effectiveness and toxicity of treatment options might be similar to the Karnofsky Performance Status scale, but applied toward treatment options instead of patient function. The conceptual framework for this scoring tool should include principles that reflect the thought processes of clinicians when evaluating treatment options. Example concepts that could be used in developing the scoring tools for the CTI are seen in Table 1.

For effectiveness, the expert panel could judge a treatment option based on concepts such as the likelihood of achieving a cure, the impact on long-term disease control, and whether it controls symptoms and improves performance status. Additionally, the level of evidence also factors into the assessment of the effectiveness of a treatment option. For toxicity, these concepts could include the probability of fatal events, severe life-threatening events, the duration of adverse effects, and the degree to which quality of life is reduced. Complicating the matter of toxicity is the fact that the degree to which a treatment's toxicity is accepted is influenced by the setting of care. For ex-

Table 1 Example Concepts Used in Developing the Comparative Therapeutic Index Scoring Tools

Effectiveness	Toxicity
<ul style="list-style-type: none"> • Level of evidence • Probability of achieving a cure • Impact on survival (e.g., overall, disease-free, progression-free) • Impact on disease control • Impact on improving performance status • Impact on disease-related symptom control 	<ul style="list-style-type: none"> • Acceptability of toxicity • Probability of fatal event • Probability of severe, life-threatening side effects • Duration of adverse effects (chronic vs. acute) • Debilitation impact of adverse effects • Impact on health-related quality of life (HR-QOL)

Comparative Effectiveness: NCCN Preliminary Report

ample, when a treatment option can potentially cure a patient (e.g., chemotherapy for testicular cancer), the toxicities, although relatively severe, may be accepted because of the potential for a dramatic, lasting benefit. Similarly, when the goal is to palliate the disease symptoms, the concerns for toxicity become increasingly important. The final toxicity scoring tool would need to incorporate this concept of the “acceptability of toxicity” based on setting of care.

Using these hypothetical concepts, the effectiveness and toxicity scoring tools could be as depicted in Tables 2 and 3. In developing these scoring tools, several principles should be applied. First, the ends of the scale (in this case 0 and 10) should be “anchored,” with a score of “0” representing no effectiveness or toxicity and a score of “10” representing the maximum theoretical effectiveness or toxicity. Maximum effectiveness could mean that a cure is readily achieved, and maximum toxicity could mean that deaths often occur. This anchoring is important to be able to establish equidistant intervals in between, and the descriptions should reflect this case. The descriptions for each interval of the scale would combine the important identified concepts, because they are not mutually exclusive and tend to overlap.

The previously identified concepts of relating effectiveness to level of evidence and having toxicity incorporate an “acceptability factor” are not concepts

that can be readily incorporated into the scale (i.e., they are not clinical outcomes). These concepts can be integrated as multiplying factors to discount the effectiveness and toxicity scores to obtain the final CTI (see Tables 4 and 5). Conceptually, the discount applied to the effectiveness score based on lower-level types of evidence takes into account that the effectiveness score may be overestimated in this situation. Applying a discount therefore results in a more conservative score. Similarly, the discount applied to the toxicity score is based on a willingness to accept some degree of toxicity based on the treatment goal.

Displaying the CTI

Given the aforementioned proposed scales for judging and rating the effectiveness and toxicity of various treatment options, the CTI should convey the aggregate scores of each treatment option rated by the panel members in a manner that is informational and user-friendly. This can be achieved by representing the effectiveness and toxicity scores of the CTI as x and y coordinates on a standard Cartesian graph (see Figure 1). In this example, the toxicity score is represented on the x axis and the effectiveness score is represented on the y axis. The general interpretation of the graph is such that treatment options clustered into the upper-left quadrant represent those with a more desirable clinical profile (with higher

Table 2 Hypothetical Effectiveness Scoring Tool

Score	Effectiveness Description
0	No meaningful impact on relevant disease-related end points
1	May provide control of symptoms and improve performance status
2	Sometimes provides short-term control of disease and may provide control of symptoms and improve performance status
3	Often provides short- to moderate-term control of disease and may provide control of symptoms and improve performance status
4	Little impact on long-term survival, often provides short- to moderate-term control of disease and may provide control of symptoms and improve performance status
5	Little impact on long-term survival, often provides moderate- to long-term control of disease, and controls symptoms and improves performance status
6	Improves long-term survival with a slight chance of cure, often provides moderate- to long-term control of disease, controls symptoms, and improves performance status
7	Improves long-term survival with a small chance of cure, often provides dramatic control of disease, controls symptoms, and improves performance status
8	Sometimes achieves a cure, improves long-term survival, often provides dramatic control of disease, controls symptoms, and improves performance status
9	Often achieves a cure, improves long-term survival, often provides dramatic control of disease, controls symptoms, and improves performance status
10	Very often achieves a cure

Li and DeMartino

Score	Effectiveness Description
0	None
1	Not known to cause death; no grade 3–4 adverse reactions; generally mild, short-term, nondebilitating toxicities and does not decrease health-related quality of life
2	Not known to cause death; low probability of grade 3–4 adverse reactions; generally mild, short-term, nondebilitating toxicities but slightly decreases health-related quality of life
3	Not known to cause death; low probability of grade 3–4 adverse reactions; generally mild, short-term, somewhat debilitating toxicities and slightly decreases health-related quality of life
4	Deaths rarely occur; low probability of grade 3–4 adverse reactions; generally moderate, short-term, somewhat debilitating toxicities and slightly decreases health-related quality of life
5	Deaths rarely occur; moderate probability of grade 3–4 adverse reactions; generally moderate, moderate-term, somewhat debilitating toxicities and moderately decreases health-related quality of life
6	Deaths rarely occur; moderate probability of grade 3–4 adverse reactions; generally moderate, long-term, somewhat debilitating toxicities and moderately decreases health-related quality of life
7	Deaths may occur; moderate probability of grade 3–4 adverse reactions; generally severe, long-term, somewhat debilitating toxicities and moderately decreases health-related quality of life
8	Deaths may occur; high probability of grade 3–4 adverse reactions; generally severe, long-term, somewhat debilitating toxicities and severely decreases health-related quality of life
9	Deaths occur; high probability of grade 3–4 adverse reactions; generally severe, long-term, debilitating toxicities and severely decreases health-related quality of life
10	Unacceptably toxic

effectiveness and lower toxicity scores) than those in the lower-right quadrant (which have lower effectiveness and higher toxicity scores). The aggregate CTI scores for each treatment option would be plotted as a 2-dimensional “scatter plot” on the graph, along with the “error bars” associated with the effectiveness and toxicity scores. The error bars of each treatment option convey the variability (e.g., standard deviation) in the judgment scoring of the panel members.

Furthermore, an electronic presentation of the CTI provides an opportunity for increased interaction with the end-user. The “points” representing the treatment options appraised by the panel could be interactive, allowing users to click on them for more information, such as the actual CTI scores for effectiveness and toxicity, the standard deviation, the NCCN category, and the setting of care. A listing of the clinically significant toxicities also could be presented, and would offer clinicians and patients

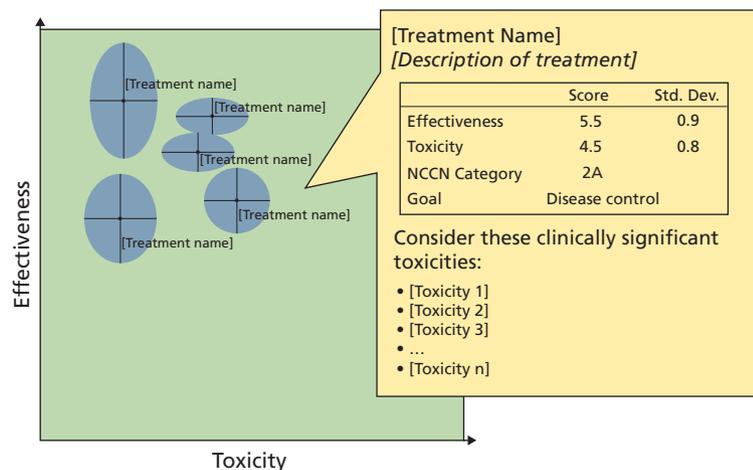


Figure 1 Proposed Display of Comparative Therapeutic Index in the NCCN Guidelines. Abbreviation: Std. Dev., standard deviation.

Comparative Effectiveness: NCCN Preliminary Report

Table 4 Hypothetical Effectiveness Score Multiplying Factor

Category of Evidence	NCCN Category 3	NCCN Category 2B	NCCN Category 2A	NCCN Category 1
Factor	0.7	0.8	0.9	1

more information to tailor treatment decisions on the individual level. In addition, a listing of the key references used to judge the CTI scores could be listed, maintaining the transparency of the process.

Scope of the CTI

The proposed CTI paradigm would be an extension of the recommendations already provided in the NCCN Guidelines (for a more detailed description of the NCCN Guidelines and the development process, please visit the NCCN Web site at www.NCCN.org). It is envisioned that the CTI scoring process will be performed during the individual guideline update meeting process. The NCCN Guideline panels meet regularly to update the guidelines based on the most current data and practice considerations, and the CTI evaluations for new and existing interventions would be added or updated accordingly. As data from actual CER studies become available, they would be integrated into the CTI clinical evaluative process.

Through providing a comparison of the relative effectiveness and toxicity based on expert judgment, NCCN is providing the end-user with additional information relevant to decision-making. However, a CTI evaluation is not appropriate for every recommendation listed in the NCCN Guidelines. Certainly, situations arise that necessitate the application of the NCCN CTI, such as those that have multiple recommendations for a particular setting (e.g., ad-

juvant chemotherapy regimens for invasive breast cancer); however, applying this paradigm may not be appropriate when very few options are available. Furthermore, when multiple treatment options are available but each will be used in a sequential manner (as in the case of stage IV, unresectable kidney cancer), the CTI does not necessarily apply. As experts, the NCCN panels will decide the most appropriate application for the CTI paradigm within their guideline.

Discussion

Clinicians, patients, and other stakeholders in oncology look to NCCN to provide leadership in the area of CER. Keeping the best interests of patients in mind, the CTI aims to be a practical, real-world, clinical evaluative process that incorporates existing data and represents the way physicians and clinicians think and practice. The objectives of the NCCN CTI are to provide additional information to clinicians and patients to better inform decision-making, with the intended goal of improving patient outcomes. To fully achieve this goal, an assessment of the challenges and unintended consequences of the CTI must be discussed.

Challenges for the CTI

Before this paradigm can be incorporated into the NCCN Guideline process, it requires proper vetting in terms of an adequate assessment of the reliability and validity of the scoring scales. Applying methodologic principles to assess these issues will require pilot testing of the CTI (i.e., the results obtained during the pilot would not be incorporated into the NCCN Guidelines). Furthermore, a working group of clinical and health outcomes experts should guide this process. Reliability must be assessed in terms of

Table 5 Hypothetical Toxicity Score Multiplying Factor

	Potentially Curative/Long-Term Control: Primary Modality	Potentially Curative/Long-Term Control: Adjuvant/Neoadjuvant Modality	Disease Control	Palliative/Salvage
Level of Acceptability Factor	0.7	0.8	0.9	1
Description	The use of this primary modality to achieve a cure or long-term disease control far outweighs toxicity	Elimination/reduction of initial, residual, or micrometastatic disease to achieve a long-term response (or potential cure) outweighs toxicity	The benefit of short- to moderate-term disease control somewhat outweighs toxicity	Improvement of disease-related symptoms or disease control is equivalent to the concerns for toxicity

whether scores can be reproduced by the same person at different intervals (intra-rater) and whether different groups of people produce similar results (inter-rater). Additionally, the validity of the scores that these scales produce must be assessed, beyond an initial “face validity” evaluation of the scoring tools (wherein clinicians are asked if the tools “look like” they will work). The assessment of validity of this type of scoring tool presents a methodologic challenge because of the nature of the information being validated. In essence, the goal is to validate whether the clinical evaluative judgment was a “good” one.

In a report that describes methods for validating these types of consensus statements, Murphy et al.²² state that it is difficult to know if a judgment is good at the time it is made. Instead, they emphasize that the focus should be on whether the process “will produce, on average, more good judgments or fewer bad judgments.” The authors mention a few methods that can be applied to the NCCN CTI process to assess whether it meets this property. These methods include an assessment of the predictive and concurrent validity. The working group of experts would decide the proper methods for this assessment, which might include comparing scores to actual practice patterns in the community and at NCCN Member Institutions. This expert working group would also modify, improve on, and reevaluate the CTI scoring system based on the reliability and validity assessments. Because this is a continual improvement process, these analyses likely will need to be performed at regular intervals to ensure a robust process.

The proposed NCCN clinically based CTI paradigm does not address an important consideration when making treatment decisions, which is the impact on resource use. In this era of focusing on the cost of health care, one can no longer ignore the impact of cost on treatment decisions. However, the “cost” of a therapy must not solely consider the acquisition price of a therapy, but ideally it should take into account other resources consumed and used during the provision of care. For example, what is the impact on hospitalization rates, use of supportive care measures (e.g., growth factors, antiemetics), and laboratory monitoring intensity? Additionally, where is the treatment administered (e.g., in the hospital, in the physician office, at home)? These factors all impact the amount of resources consumed and relates to an episodic cost of care.

Clinicians may take into consideration these matters when tailoring therapy to a particular patient. It is recognized that the proposed CTI does not address these issues, because the focus is narrowed toward the clinical implications of various treatment options. As they do with clinical issues, however, others look toward NCCN for leadership and guidance in the area of health economics of cancer care. Therefore, guidance that NCCN provides in this area would likely be a separate resource, outside of the NCCN Guidelines process, and a derivative product thereof.

Future Direction

As experts in the dissemination of evidence-based recommendations, NCCN is well positioned to actively disseminate CE results and the clinical evaluative judgments made based on these reports. The NCCN Web site, www.NCCN.org, attracts more than 1.3 million unique visitors per year. Beyond the NCCN Guidelines, NCCN has many other programs and resources to inform and improve decision-making and outcomes for patients. NCCN’s spectrum of programs and resources emphasizes improving the quality, effectiveness, and efficiency of oncology practice. NCCN hosts educational conferences and symposia for physicians at which the CTI and its results from different panels and settings can be presented to encourage change in practice patterns so that patients receive the best, appropriate care.

Ultimately, the future direction of NCCN’s CER initiative could be to develop a continuous learning system, as initially described by the Friends of Cancer Research, whereby the most current and best evidence would be translated into practice recommendations, with a subsequent evaluation of outcomes based on these recommendations.⁷ In considering the application of the CTI for clinical evaluative comparisons of treatment options, NCCN is uniquely positioned to assume a leadership role in oncology CER, especially with this type of translation and adoption of CER. NCCN is recognized in oncology as the arbiter of high-quality cancer care based on NCCN world-leading institutions and clinicians, and the status of the NCCN Guidelines as the standard for clinical policy in oncology in the United States.

In moving forward, NCCN will reconvene its CE Work Group that includes representatives from the patient community, academia, community practice, payor community, and the pharmaceutical/

Comparative Effectiveness: NCCN Preliminary Report

biotech industry. NCCN will work to validate the evaluative scoring systems and assess the feasibility of integrating the proposed CTI process into the efforts of the NCCN Guideline panels. Only after these important issues are duly considered and studied will a final decision be made about establishing the proposed CTI process as a tool for decision-making. In the end, NCCN seeks to enhance the value of the current recommendations in NCCN Guidelines by introducing a more direct comparative component to facilitate and improve decisions on behalf of patients.

References

- Centers for Medicare and Medicaid Services. Projected National Health Expenditure Data. Available at: http://www.cms.hhs.gov/NationalHealthExpendData/03_NationalHealthAccountsProjected.asp. Accessed October 13, 2009.
- Congressional Budget Office. Research on the comparative effectiveness of medical treatments. December 2007. Available at: <http://www.cbo.gov/ftpdocs/88xx/doc8891/12-18-ComparativeEffectiveness.pdf>. Accessed July 27, 2010.
- Congressional Budget Office. Health care. Available at: <http://www.cbo.gov/publications/collections/collections.cfm?collect=10>. Accessed June 21, 2010.
- Institute of Medicine National Research Council. Ensuring Quality Cancer Care. Washington, DC: National Academy Press; 1999.
- United States. Cong. Senate. American Recovery and Reinvestment Act of 2009. 111th Cong., 1st sess. Washington: GPO, 2009.
- Kaiser Family Foundation. Summary of New Health Reform Law. Available at: <http://www.kff.org/healthreform/upload/8061.pdf>. Accessed June 15, 2010.
- Friends of Cancer Research. Improving medical decisions through comparative effectiveness research: cancer as a case study. Available at: http://focr.org/files/CER_REPORT_FINAL.pdf. Accessed October 9, 2009.
- Shumock GT, Pickard AS. Comparative effectiveness research: relevance and applications to pharmacy. *Am J Health Syst Pharm* 2009;66:1278–1286.
- Tunis SR, Stryer DB, Clancy CM. Practical clinical trials: increasing the value of clinical research for decision making in clinical and health policy. *JAMA* 2003;290:1624–1632.
- Teutsch SM, Berger ML, Weinstein MC. Comparative effectiveness: asking the right questions, choosing the right method. *Health Affairs* 2005;24:128–132.
- Berger ML, Mamdani M, Atkins D, et al. Good research practices for comparative effectiveness research: defining, reporting and interpreting nonrandomized studies of treatment effects using secondary data sources: the ISPOR Good Research Practices for Retrospective Database Analysis Task Force Report – Part I. *Value Health*. Epub 2009 Sep 29.
- Cox E, Martin BC, Van Staa T, et al. Good research practices for comparative effectiveness research: approaches to mitigate bias and confounding in the design of nonrandomized studies of treatment effects using secondary data sources: the International Society for Pharmacoeconomics and Outcomes Research Good Research Practices for Retrospective Database Analysis Task Force Report – Part II. *Value Health*. Epub 2009 Sep 10.
- Johnson ML, Crown W, Martin BC, et al. Good research practices for comparative effectiveness research: analytic methods to improve causal inference from nonrandomized studies of treatment effects using secondary data sources: the ISPOR Good Research Practices for Retrospective Database Analysis Task Force Report – Part III. *Value Health*. Epub 2009 Sep 29.
- Lee SJ, Earle CC, Weeks JC. Outcomes research in oncology: history, conceptual framework, and trends in the literature. *J Natl Cancer Inst* 2000;92:195–204.
- Inglehart JK. Prioritizing comparative-effectiveness research—IOM recommendations. *N Engl J Med* 2009;361:325–327.
- Naik AD, Petersen LA. The neglected purpose of comparative-effectiveness research. *N Engl J Med* 2009;360:1929–1931.
- Garber AM, Tunis SR. Does comparative-effectiveness research threaten personalized medicine? *N Engl J Med* 2009;360:1925–1927.
- Rebeck TR. Inherited genetic markers and cancer outcomes: personalized medicine in the postgenome era. *J Clin Oncol* 2006;24:1972–1974.
- Department of Health and Human Services. Draft Definition, Prioritization Criteria, and Strategic Framework for Public Comment. Available at: <http://www.hhs.gov/recovery/programs/cer/draftdefinition.html>. Accessed October 2, 2009.
- Sox HC, Greenfield S. Comparative effectiveness research: a report from the Institute of Medicine. *Ann Intern Med* 2009;151:203–205.
- U.S. Department of Health and Human Services FDA Center for Drug Evaluation and Research; U.S. Department of Health and Human Services FDA Center for Biologics Evaluation and Research; U.S. Department of Health and Human Services FDA Center for Devices and Radiological Health. Guidance for industry: patient-reported outcome measures: use in medical product development to support labeling claims: draft guidance. *Health Qual Life Outcomes* 2006;4:79.
- Murphy MK, Black NA, Lamping DL, et al. Consensus development methods, and their use in clinical guideline development. *Health Technol Assessment* 1998;2:21–22.