Overview: Comparative effectiveness research (CER) attempts to compare the benefits and harms of alternative strategies for diagnosing, treating, or preventing a disease in typical patients in the population. Despite enormous and rapidly rising health care expenditures in the United States, outcomes considered valid measures of health have not improved in a corresponding fashion. The goal of CER is to define the optimal strategies and target population for delivering the most effective and safe interventions in the most efficient way. Cancer has been identified as one of the high priority areas for such research and is the focus of a substantial influx of additional federal research dollars. Although randomized controlled trials and meta-analyses of such trials have and continue to represent the gold standard for comparative effectiveness studies, the limitations of randomized trials necessitate consideration of a broad range of research approaches including cohort, population, and modeling studies, among others. Although such investigations have their own very important limitations, the goal of CER in oncology is to gather the totality of available evidence to address critical questions related to the comparative benefit, harm, and overall value of established and emerging interventions. CER is often confused with restrictions, rationing, and cost containment issues. However, CER appropriately guided from an objective yet knowledgeable clinical perspective attempts to distinguish what works in typical patients from what does not. Such studies should strive to objectively compare available options so that patients, providers, and policymakers can make rational, clinically sound, and evidence-based choices to improve the health care of both individuals and society.

The Institute of Medicine (IOM) has provided a working definition of CER as the “generation and synthesis of evidence that compares the benefits and harms of alternative methods to prevent, diagnose, treat and monitor a clinical condition, or to improve the delivery of care. The purpose of CER is to assist consumers, clinicians, purchasers, and policy makers to make informed decisions that will improve health care at both the individual and population levels.” Fundamental to this definition are the direct comparisons of the effectiveness and safety of different interventions and investigations in “typical” patients seen in everyday practice. More than $1 billion has been appropriated for CER through the Department of Health and Human Services ($400 million), the National Institutes of Health ($400 million), and the Agency for Healthcare Research and Quality ($300 million). The primary goal of this large appropriation is to begin to address the challenges and provide innovative strategies for the evaluation, approval, and utilization of existing and new technologies. The American Recovery and Reinvestment Act of 2009 called on the IOM to recommend a list of priority topics as the initial focus of this investment in CER.

Although cancer was only included in a limited fashion in seven of the top 100 IOM priorities, nowhere is the need and potential for CER greater than in the area of oncology. Advances in molecular biology and genetics have led to the identification of countless potential targets for diagnostic, prognostic, and predictive assays as well as targets for the treatment and prevention of cancer. Enthusiasm surrounding the rapid and accelerating pace of new diagnostic and therapeutic innovations is muted, however, by the continuing relentless increase in health care costs threatening the very economic stability of the country. Even more distressing is evidence that outcomes achieved from that health care spending are poor in relationship to per capita health expenditure in the United States.

To address this challenging environment, it is essential to identify important diagnostic and therapeutic questions and to acquire, evaluate, and, where appropriate, synthesize all relevant information related to effectiveness, safety, and overall value of comparative diagnostic, treatment, and prevention strategies in order to properly inform important clinical and policy decisions.

Randomized Controlled Clinical Trials

Comparative results from prospective randomized controlled clinical trials (RCTs) or meta-analyses of such trials represent the gold standard of CER. Importantly, well-designed and conducted RCTs assure the investigator that both recognized and unknown factors that might influence the outcome of interest will be equally distributed on average between the study arms and should not influence treatment effect. Unfortunately, not all RCTs in oncology are of good quality as some trials lack adequate sample size, are not properly conducted or designed, or are not appropriately analyzed. Systematic reviews and evidence synthesis in the form of meta-analyses of RCTs should define a priori a rigorous process of search, review, selection, data abstraction, and appropriate analytic methodology including an assessment of heterogeneity across studies as well as study quality (Fig. 1).

However, most clinical questions in oncology have not been the subject of RCTs. Although placebo-controlled RCTs are essential to properly address supportive care questions, they are often considered problematic or even unethical in comparative studies of cancer treatment. Even when completed, many methodologic challenges remain, including the use of short-term or surrogate efficacy outcomes, limited sample size, or duration of follow-up to address important toxicity or safety questions, and often only passing attention to patient-reported outcomes reflecting the comparative effect of treatment on quality of life. Finally, RCTs frequently use strict and limiting eligibility criteria that exclude patients with major but common comorbidities. These comorbidities may be encountered in older patients with cancer,
may reduce effectiveness or increase toxicity, and often disqualify such patients from clinical trials. Therefore, RCTs often do not adequately address effectiveness in the broader, unselected cancer population with major medical comorbidities and treatment safety issues that may not emerge until years later. Therefore, alternative sources of evidence are needed to guide the evaluation and approval of new interventions and address the immediate and compelling need for patients, clinicians, and policymakers to make critical clinical and policy decisions in the absence of generalizable data from multiple or large well-designed RCTs.

**KEY POINTS**

- Health care costs continue to escalate rapidly without a corresponding increase in the overall health of the population.
- Comparative effectiveness research (CER) aims to compare available options for the diagnosis, treatment, and prevention of disease in typical patients.
- CER in oncology seeks to contrast the benefits, harms, and overall value of available conventional and emerging interventions.
- Although randomized controlled trials remain a mainstay of comparative effectiveness research, other study designs including cohort, population, and modeling approaches may contribute important information for comparative studies.
- Major challenges for CER include the appropriate integration of a clinical perspective into the design of such studies, as well as the translation of study results into clinical practice and public policy.

**Other Sources of Evidence on Comparative Effectiveness**

The limitations of available evidence from RCTs have led investigators to explore additional sources that, when properly applied, may provide reasonable, valid, and perhaps more generalizable estimates of comparative effectiveness, safety, and costs and may also generate hypotheses that form the basis of future confirmatory RCTs (Table 1). The challenge of CER when the intervention of interest is not randomly assigned is to adjust for possible confounding factors that are associated with both the outcome and the intervention and that may account for or obscure the treatment effect. It is essential that the same rigorous attention to study design, conduct, analysis, and reporting be applied to comparative observational studies as are accepted measures of high-quality RCTs. The study hypothesis, study population, controls, measurements, analytic methods, and any subgroup analyses should be defined a priori in a written protocol approved by an investigational review board. However, there remain many challenges to studies based on data not specifically designed for comparative studies but rather to capture administrative, claims, or tumor registry data. Such studies may be limited by inaccurately recorded clinical measures, a large number of missing variables, and the choice of an intervention at the discretion of the treating clinician, which may be influenced by

**Table 1. Toolbox for Comparative Effectiveness Research**

<table>
<thead>
<tr>
<th>Study Design</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Randomized controlled trials</td>
<td></td>
</tr>
<tr>
<td>Systematic reviews and meta-analyses</td>
<td></td>
</tr>
<tr>
<td>Other comparative clinical trials</td>
<td></td>
</tr>
<tr>
<td>Population studies including registries and administrative and claims data</td>
<td></td>
</tr>
<tr>
<td>Prognostic and predictive association studies</td>
<td></td>
</tr>
<tr>
<td>Quality-of-life studies including patient reported outcomes</td>
<td></td>
</tr>
<tr>
<td>Clinical decision models including cost effectiveness and cost utility analyses</td>
<td></td>
</tr>
</tbody>
</table>

Reproduced with permission from Lyman.15
clinical factors that also affect prognosis and response to the treatment.

Other approaches used in CER include complex but carefully constructed and populated clinical simulations. Taking advantage of the totality of available evidence from RCTs, meta-analyses, and population studies along with the experience and judgment of expert clinicians, clinical decision models attempt to emulate realistic clinical scenarios. The main advantage of such models is their ability to address specific and relevant clinical questions applying the best available data while accounting for measurement variability or uncertainty permitting reasonable estimates of the comparative effectiveness, safety, and cost of an intervention. Clinical decision simulation permits the investigator to vary assumptions about the decision choices, probabilities of events, and value of important outcomes to study the influence of these assumptions on the best decision. Comparative effectiveness may be assessed when clinical conditions, including the values of assumptions, change over time using a Markov modeling approach. Such simulations have been applied to important clinical questions for the purpose of modeling the comparative or incremental effectiveness, utilities reflecting the effect on quality of life, costs, and composite measures such as cost-effectiveness and cost utility (i.e., the cost per year of life or quality-adjusted year of life saved). Such modeling has been applied to a wide range of conventional as well as novel oncologic interventions with interesting results (Table 2). The obvious policy implication of CER is that, where the incremental benefit is small and the expenditures are large, the limited value of some interventions might be more effectively applied to more promising and valued clinical strategies.

Comparative Effectiveness of Personalized Medicine Approaches

The potential for providing more targeted individualized or personalized interventions, tailoring treatments to patients at greatest risk and most likely to benefit has garnered considerable enthusiasm among researchers, clinicians, and patients. As noted above, the highest level of evidence for the comparative effectiveness studies of personalized approaches is through large, well-designed, prospective RCTs or meta-analyses of such trials comparing guided or targeted interventions with concurrent unguided interventions. Other comparative effectiveness studies with well-matched concurrent control subjects but otherwise rigorous methodology may sometimes be used (Table 3). Such true comparative effectiveness studies must be distinguished from association studies utilized in the development (phase I) or validation (phase II) of prognostic or predictive models. Although such studies can only indirectly address the comparative effectiveness and value of such approaches, they are critical to the development of innovative strategies aimed at a goal of personalized medicine. Retrospective analyses of prospective clinical trials or cohorts with valid outcome evaluation along with archived samples permitting assessment of relevant genetic and molecular markers have been successfully pursued.

Perhaps the greatest opportunity as well as challenge for CER in oncology relates to the evaluation of personalized approaches based on molecular and genetic biomarkers for identifying patients at risk for disease, for disease recurrence, or for experiencing benefit or harm with new interventions. Targeted personalized interventions have the potential to increase the effectiveness of existing as well as new agents while reducing treatment-related complications by identifying patients at high or low risk. Since treatment-related complications represent a major cause of morbidity and mortality in the general cancer population, targeted interventions are designed to minimize toxicities while maximizing treatment effectiveness and cost-effectiveness. More selective and targeted use of new expensive technologies may also reduce health care expenditures or maximize the effectiveness gained with the resources applied in the care of patients with cancer.

The role of genetic and molecular biomarkers must be rigorously evaluated like any diagnostic, prognostic, or predictive test considering both assay performance as well as the patient population studied. A systematic review of studies of gene profile assays to predict disease recurrence in patients with early-stage breast cancer demonstrated considerable variation in predictive assay performance. To assess the comparative effectiveness and potential value of such assays, clinical decision simulation studies have been conducted based on input from clinical experts and data from the clinical trials used to evaluate the prognostic and predictive validity of the assay. Recommendations for the standardized complete and transparent reporting of prognostic and predictive biomarker studies in oncology have been proposed.

Biomarker assays will eventually be developed for many novel agents providing opportunities for better targeting toward patients most likely to benefit while avoiding exposure of patients unlikely to respond. However, the development of targeted therapies in advance of the corresponding predictive assay for the potential molecular target represents an emerging challenge for CER in oncology. This is
perhaps best illustrated by the recent development of inhibitors of the epidermal growth factor receptor and assays identifying mutations of the KRAS oncogene, assays for which were only validated during or following completion of the pivotal trials.\textsuperscript{14} It is important to anticipate such assays when conducting large confirmatory trials for targeted agents by archiving tissue and blood for future validation of prognostic and predictive biomarkers. Recommendations have been provided for the rigorous codevelopment of biomarkers in trials of targeted treatments to avoid clinical or regulatory dependence on exploratory studies (Table 4).\textsuperscript{15} Pivotal trials of targeted biologic therapies should be encouraged to archive samples on all subjects in order to aid rapid validation of new biomarkers in the future. It will also be essential for assessment of comparative effectiveness and safety that individual patient data be shared across trials for pooled analyses of biomarker performance.\textsuperscript{16}

Public Policy and Clinical Practice Guidelines

Although CER often leads to the generation of clinical practice guideline recommendations as well as changes in health care policy and reimbursement, it is important to note that those actions are not inherent to such research but result from the interpretation of the results of such research as well as other available evidence and opinion. An example of where the line of separation between CER and guideline recommendations was blurred is the recent controversy over updated recommendations on Breast Cancer Screening from the U.S. Preventive Services Task Force.\textsuperscript{17} After commissioning both an updated systematic review and meta-analysis and a sophisticated decision modeling effort, the panel came to quite different recommendations compared with previous guidelines specifically related to the age for initiating routine breast cancer screening, the appropriate screening interval, and the value of breast self-examination.\textsuperscript{18,19} The updated systematic review added one additional RCT and updated two others estimating an overall relative risk reduction in breast cancer mortality of 15% in 40- to 50-year-old women who were randomly selected for breast cancer screening across the eight RCTs. Despite these results and a nearly 25% reduction in breast cancer mortality observed since Medicare approved coverage for screening mammography in the early 1990s, the panel recommended against routine screening mammography in women age 40 to 49 and for biennial rather than annual screening starting at age 50. In issuing the new guidelines, the panel raised concerns about the risk of false-positive mammograms leading to additional imaging or biopsies. The change in recommendations from the panel based on their interpretation of the evidence may have been justifiably criticized. However, it should not be equated with the CER that went into the evidence summary presented to the panel. Guideline panels for the American Society of Clinical Oncology also perform a systematic review of the available evidence and then make recommendations based on review and interpretation of the data in the context of good clinical experience and judgment from both methodology and disease content experts.\textsuperscript{4}

The Importance of Clinical Comparative Effectiveness Research

CER should not be considered a backdoor method for avoiding the effort and cost of conducting large, well-designed RCTs or meta-analyses of such trials, which remain the gold standard of comparative effectiveness. In fact, RCTs represent an excellent source of information to inform subsequent population and decision-modeling studies. However, in the absence of data or with conflicting results from high-quality RCTs or where such trials are deemed unethical or not feasible, there are many additional research tools available to consider in evaluating comparative effectiveness to enhance evidence from RCTs. CER should not represent an avenue to bypass RCTs to achieve some desired policy objective, nor should it be a strategy for restricting reimbursement for effective but costly treatments. Rather, CER is an important mechanism for systematically identifying and summarizing the totality of evidence on the effectiveness, safety, and value of competing strategies including results from RCTs and meta-analyses, population studies, and clinical simulation studies in order to inform and provide patients, providers, and policymakers with valid recommendations on the optimal management of patients with cancer.

The current discussion around CER should be recast as comparative “clinical” effectiveness research recognizing the importance of the clinical setting, key clinical issues to be addressed, clinically relevant study designs, use of relevant clinical outcomes including measures of effectiveness and toxicity, and appropriate clinical interpretation of results allowing reasonable integration of research findings into clinical practice. Although always cognizant and guarding against potential conflicts of interest, comprehensive integration of the clinical perspective into the process will enhance the relevance and utility of CER. Likewise, it will reduce the chance of arriving at unwarranted conclusions that may compromise patient care and waste enormous resources addressing irrelevant questions and inappropriate study designs. CER, with appropriate clinical guidance, provides the greatest opportunity for identifying the most effective and safe interventions for the optimal management of patients with cancer while defining the comparative value of the diagnostic and therapeutic options available for improving cancer patient care.
Author’s Disclosures of Potential Conflicts of Interest

<table>
<thead>
<tr>
<th>Author</th>
<th>Employment or Leadership Positions (Commercial Firms)</th>
<th>Consultant or Advisory Role</th>
<th>Stock Ownership</th>
<th>Honoraria</th>
<th>Research Funding</th>
<th>Expert Testimony</th>
<th>Other Remuneration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gary H. Lyman</td>
<td></td>
<td>Amgen</td>
<td></td>
<td></td>
<td>Research Support to Duke University (PI)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

REFERENCES