Comparative effectiveness research:

What it means, what it means for hospitals

Lawton R. Burns

Lawton R. Burns, Ph.D., MBA, is the chair of the Health Care Management Department, and also the James Joo-Jin Kim Professor, as well as a professor of healthcare systems at the Wharton School at the University of Pennsylvania. He is also director of the Wharton Center for Health Management and Economics. Dr. Burns has served on the Governing Board of the Institute of Medicine (Health Services section) and serves on the editorial board of Health Services Research. He is a past member of the Grant Review Study Section for the Agency for Health Care Policy and Research. He is also a Life Fellow of Clare Hall at the University of Cambridge.

Evolution of comparative effectiveness research

Comparative effectiveness research (CER) has evolved gradually over time. CER has its roots in randomized controlled trials (initiated in the 1940s), technology assessment (1970s), outcomes research (1980s), evidence-based medicine (1990s), the Effective Health Care Program (Section 1013 of the 2003 Medicare Modernization Act, developed in 2005), the American Recovery and Reinvestment Act (2009), and most recently in the Patient Protection and

Affordable Care Act (2010). All of these historical developments have contributed to what we now understand as CER.

What is comparative effectiveness research?

Varying definitions of CER have been offered over the past few years by the Congressional Budget Office, the National Institutes of Health, the Medicare Payment Advisory Commission, the Agency for Healthcare Research and Quality, the Institute of Medicine, and the American College of Physicians. At the heart of CER is the conduct and synthesis of research that compares the benefits and costs of different treatment interventions for a given medical condition. This goes beyond the normal randomized controlled trial of a new drug versus a placebo. CER uses primary and secondary data to evaluate the relative effectiveness of two or more therapeutic treatments – e.g., a drug versus another drug, a device versus another device, a drug versus a device, a drug (device) versus a surgical or medical procedure, etc. – in realistic healthcare settings. Relative effectiveness is measured in terms of the differences in patient outcomes (morbidity, mortality, adverse events, quality of life, symptoms) and/or the differences in cost of the various interventions considered.

For those interested, CER involves more than just research, however. CER also entails investments in human and scientific capital (training), data infrastructure (e.g., using electronic health records), and dissemination and translation of the results. Such investments need to be targeted at priority patient populations, medical/surgical conditions, and types of interventions. Interventions can go beyond treatments and products to also include behavioral changes, delivery system changes and prevention strategies. While drugs and devices require approval for market entry, procedures and behavioral changes and delivery system changes do not.

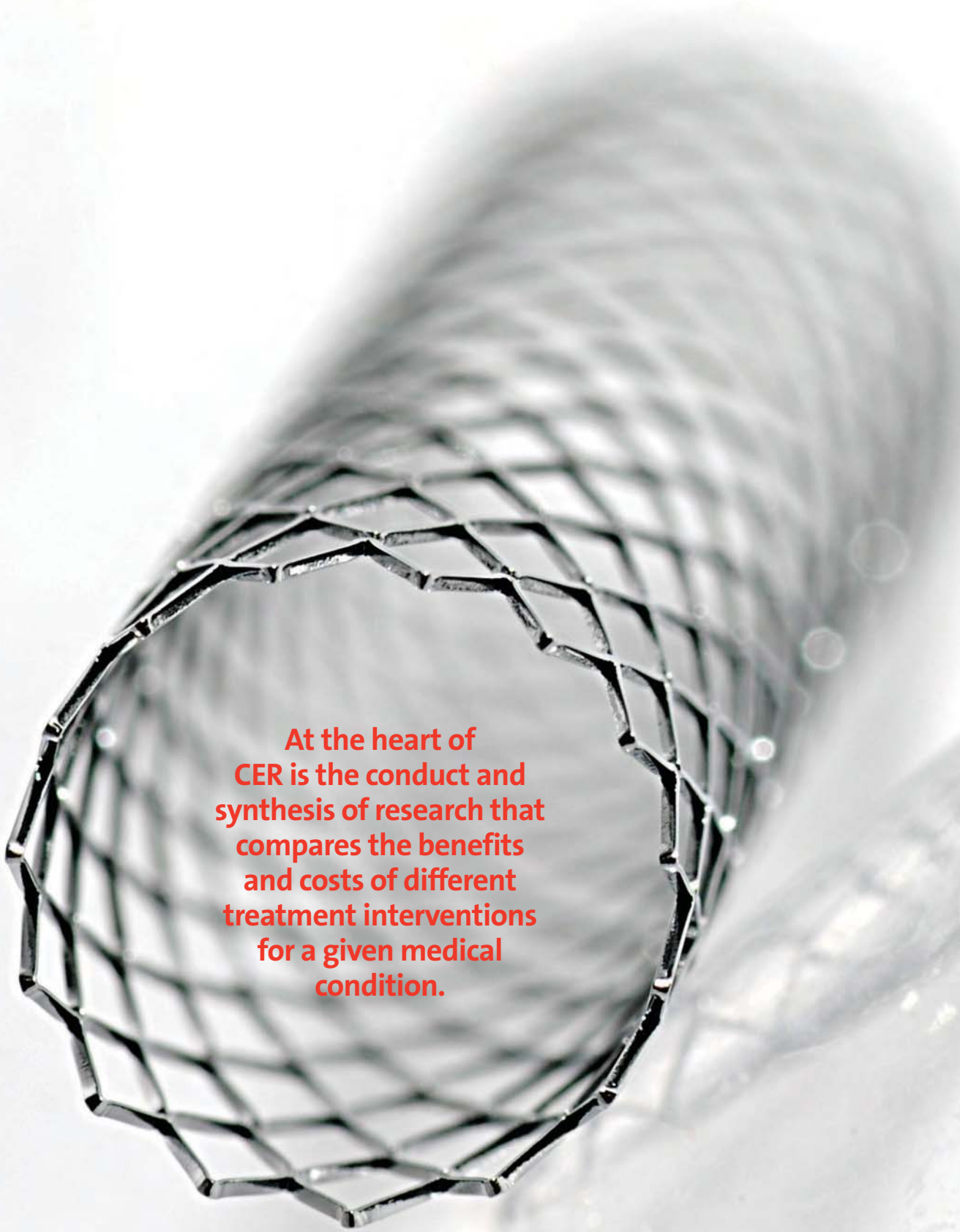
What is the overarching purpose that CER is designed to serve? CER is intended to help public and private sector payors "bend the trend" in healthcare costs – whether by informing supply chain management purchasing, reducing geographic variations, promoting evidence-based medicine, or increasing value-based purchasing. CER also aims to promote higher quality of care in addition to lower costs, and thereby increase the value of healthcare provided.

Study designs in CER and relevance for physician preference items (PPIs)

CER studies can be undertaken in a multitude of formats. They can include systematic reviews of the literature (including meta-analyses), decision modeling studies, retrospective analyses of clinical data, prospective experimental and non-experimental studies, and pragmatic prospective head-to-head trials. The study formats involve different trade-offs in internal validity, generalizability, feasibility, time and cost.

CER research is seldom conducted on PPIs. For example, there are a limited number of head-to-head studies of (1) alternative drugs – e.g., the Clinical Antipsychotic Trials of Intervention Effectiveness (CATIE) for antipsychotic medications, the Antihypertensive and Lipid-Lowering Treatment to Prevent Heart Attack (ALLHAT) trial of antihypertensive and lipid lowering medications; and (2) alternative medical devices and procedures – e.g., the Clinical Outcomes Utilizing Revascularization and Aggressive Drug Evaluation (COURAGE) trial comparing stains with PCI, the COMPARE trial evaluating different drug-eluting stents.[1]

Why so little effort and attention here? There are several general reasons and some reasons specific to devices. First, large manufacturers of incumbent products may resist CER since they stand losing market share if trial results are not favorable. Conversely, small manufacturers of potentially disruptive technologies typically lack the financial resources needed to conduct the trials. Second, CER trials are expensive and require a lot of patients using a lot of different products to detect important outcome differences among them, not to mention the time



At the heart of CER is the conduct and synthesis of research that compares the benefits and costs of different treatment interventions for a given medical condition.

needed for patient enrollment and product evaluation. Third, CER studies sponsored by manufacturers are likely tainted by suspicions of bias and conflicts of interest, and hence might not be well- regarded by clinicians. With regard to medical devices specifically, new devices are evaluated primarily in terms of safety, not efficacy. Most new device approvals take the 510(k) premarket notification route (which requires only demonstrating substantial equivalence to another device already on the market) rather than the premarket approval route (which requires trials to demonstrate clinical safety and efficacy), and thus are subject to faster and less rigorous approval.[2] The outcomes for some devices (e.g., hip implants) may not be observable in the short term, but rather require long-term patient follow-up. Finally, given the short product lifecycles in medical devices (18-24 months), CER findings are often outdated by the time they are published due to the release of a new generation product.

Concerns with CER

In addition to the time and cost issues, CER faces other hurdles and concerns. Any new product or procedure is likely to be effective for some patient populations. Very few are likely to be totally useless for everyone. It is well-known that most widely prescribed drugs have efficacy for perhaps half of the patients taking them. This means there are no black-and-white answers for a given product. Any product evaluations will need to be conditional (e.g., value-added for certain patients but not all).

Moreover, some of the products and therapies to be compared are at the heart of ongoing turf wars between specialists. Thus, interventional cardiologists who implant drug-eluting stents compete with cardiac surgeons performing coronary artery bypass with graft. Recent vendor-sponsored trials may show that the former outperforms the latter. Institutions that conduct CER need to be cognizant of the interests of competing specialists who use the technologies being evaluated, as well as their brand preferences.

Two additional concerns are the cost-effectiveness of CER and its track record. A 2008 report by the Congressional Budget Office suggested that CER investments over time outweigh the modest savings in healthcare costs borne by the government. Studies that have analyzed the impact of CER in drug trials (e.g., CATIE, COURAGE) have found little impact on physician prescribing patterns. This is due to the lack of consistency between the study findings and physicians' clinical practice, physician skepticism about the CER results, and the absence of significant incentives to get physicians to change practice patterns. Indeed, implementation of CER results may be a murky black box: how to translate research evidence into practice. There is some evidence that clinicians do not always follow practice guidelines or prompts from clinical decision support systems for a host of reasons that are only vaguely understood.

CER advocates also need to recognize there are *multiple* determinants of treatment effectiveness. In addition to the choice of treatment option, treatments vary in effectiveness based on individual patient characteristics and delivery system characteristics. Right now researchers do not know the relative importance of these three sets of factors. Any CER study that is undertaken should endeavor to compare not only the relevant treatment options, but also to capture as much information as possible on the patient and the delivery setting. With the push for personalized medicine, there will be a need to study how treatments work in patient sub-populations and respond differently in different patient genotypes.

CER will also require some major balancing acts among the stakeholders in the healthcare system. As one example, payors want value (quality divided by cost) for their money; manufacturers, on the other hand, want money for the value they argue their products render. CER is necessary but not sufficient to satisfy both of these objectives. As another example, the U.S. has experienced an unrelenting rise in national health expenditures over the last 60-70 years (roughly 2-3 percent above inflation). New technological advances and their widespread application to the population are the biggest drivers of these costs.[3] Our nation has yet to figure out how to balance the seemingly limitless innovation in products and therapies with the increasingly constrained public and private payor budgets – other than letting providers at the bedside handle the rationing. CER may help providers in this difficult endeavor. Finally, CER is designed to promote value in purchasing and yet should not dampen innovation. CER studies can have product winners and losers. Losing out in a CER study may retard innovation and entrepreneurship, may limit competition among vendors, and may inhibit diffusion of products that later prove valuable.

Likely impact of CER on the industry

Some general observations: CER will have two immediate effects. First, there will be heightened scrutiny of new medical technology, especially given its role in rising healthcare costs and the heretofore lack of requirement for data on product efficacy for approval. Payors and providers will increasingly gather comparative data on medical products and services that are used. Second, this scrutiny will issue forth in more research activity, more employment of CER-trained researchers, more monies budgeted for CER studies, and more scrutiny of the practices of physicians who employ these technologies.

There will be other effects as well. We are already seeing an effort to conduct new types of research (population health studies) using new types of databases (e.g., patient registries), such as Kaiser's effort to monitor total joint replacements using its orthopedic registry. We will likely see more research conducted outside of academic settings in community hospitals and other 'real world' settings. We may also see rising demand for products deemed "winners" in these studies and falling demand for those deemed "losers," although prior experience suggests that

such market share shifts are slow to develop due to physician reluctance to change practice patterns.

Impact on suppliers: To prepare for this brave new world of CER, product manufacturers will likely devote more in-house attention to comparative research on their own products and those of their rivals. Rather than let outside researchers develop such comparative evidence, vendors may seek to do this themselves (or at least have their own results with which to champion their products). The new "risk mitigation strategy" will consist of auditing the evidence for one's products and those of competitors, monitoring what products are being targeted by payors for technology assessment and CER reviews, designing more rigorous research protocols, launching one's own CER studies, gathering more outcomes data and data on more types of outcomes (e.g., safety, efficacy, quality of life, productivity, costs), and identifying patient sub-populations in which one's products work particularly well. This will have the effect of diverting some monies that would have gone to sales and marketing and re-allocating them to additional R&D, thereby lengthening product development times.[4] The CER scrutiny will certainly impel suppliers to clinically differentiate their products in verifiable ways, since CER findings have the potential to commoditize current PPIs. This may raise the innovation threshold for new projects undertaken by some suppliers and reduce incremental innovation projects. This is exacerbated by current changes in the regulatory environment whereby the FDA is raising requirements for 510(k) approvals. All these changes may disadvantage smaller startups (and the venture capitalists who fund them).

Impact on hospitals: Hospitals and other healthcare providers will likewise engage in more CER studies and develop/harness the information technologies needed to form the CER data infrastructure. These analyses can be applied to both products as well as to hospital operations and the design of the delivery system. With regard to products, CER data can inform the deliberations of hospital value analysis teams who evaluate new products seeking to enter the hospital. The CER data can enrich the conversations between materials managers and the clinicians championing the use of particular PPIs, and thereby broaden clinicians' participation in supply chain management activities. CER data can also be used to increase the hospital's bargaining leverage with suppliers (both on the pharmacy side as well as materials management side). Hospital systems and group purchasing organizations can partner on product evaluations and trials, leveraging data across all member hospitals to compare product costs versus procedure and operating costs to identify sources of efficiency, and conducting head-to-head prospective studies. As one illustration, Wharton School researchers gathered and evaluated surgeons' assessments of the ergonomic performance of suture and endomechanical products made by eight different vendors in animal labs housed in academic medical centers around the country. Surgeons were recruited from the member hospitals to participate in the trials. Head-to-head comparisons documented the superior performance of one vendor over the others, with two additional vendors coming in second place at an acceptable level.[5] With regard to hospital operations, CER might be applied to treatment delivery failures in such areas as fall prevention, patient flow management, computerized physician order entry systems, duplicate reading of imaging studies, and pharmacists' participation on clinical rounds.

CER lessons

The application of CER in the past has met with only mixed success. What needs to change going forward? Evidence on products and therapies needs to be gathered more frequently (i.e., across a variety of drugs, devices, procedures, etc.) and more quickly prior to widespread adoption to inform clinician decision-making. To implement CER, providers and payors will need to think about the incentives that physicians need in order to change their practice patterns. The topics selected for CER analysis should address major areas of cost in hospitals such as PPIs. Finally, differences in the sub-populations receiving the treatment (patient characteristics) and the context in which the treatments are used (delivery system characteristics) need to be measured to capture all of the drivers of treatment effectiveness.

CER that is "done right" will have timely evidence that is relevant to providers, patients, payors and vendors. It should address issues of "overuse," "underuse," and "misuse" of medical technology, all of which can harm quality of care and increase costs. Finally, CER should serve to enhance the practice of evidence-based medicine.

References:
1 McEvoy JP, Lieberman JA, Stroup TS, Davis SM, Rosenheck RA, Swartz MS, Perkins DO, Keefe RSE, Davis CE, Severe J, Hsaio JK for Catie Investigators. "Effectiveness of Clozapine, Olanzapine, Quetiapine and Risperidone in Patients with Chronic Schizophrenia Who Failed Prior Atypical Antipsychotic Treatment." American Journal of Psychiatry. 163:600-610 (2006). ALLHAT Officers and Coordinators for the ALLHAT Collaborative Research Group. "Major Outcomes in High-Risk Hypertensive Patients Randomized to Angiotensin-Converting Enzyme Inhibitor or Calcium Channel Blocker vs Diuretic." Journal of American Medical Association 288: 2981-2997 (2002). Boden WE, O'Rourke RA, Teo KK, et al. "Optimal Medical Therapy with or without PCI for Stable Coronary Disease." New England Journal of Medicine 356:1503-1516 (2007). Kedhi E, Joesoef KS, McFadden E, et al. "Second-Generation Everolimus-Eluting and Paclitaxel-Eluting Stents in Real-Life Practice (COMPARE): A Randomised Trial." Lancet 375: 201-209 (2010).
2 Kruger K. "The Medical Device Sector." In Burns LR (Editor), The Business of Healthcare Innovation (Cambridge, UK: Cambridge University Press), Chapter 6 (2005).
3 Burns LR. "The Business of Healthcare Innovation in the Wharton School Curriculum." In Burns LR (Editor), The Business of Healthcare Innovation (Cambridge, UK: Cambridge University Press), Chapter 1 (2005).
4 Kruger. "The Medical Device Sector."In Burns LR (Editor), The Business ofHealthcare Innovation (Cambridge,UK: CambridgeUniversity Press), Chapter 6 (2005).
5 Burns LR, Lee JA, Bradlow ET, Antonacci AC. "Surgeon Evaluation of Suture and Endo-Mechanical Products." Journal of Surgical Research. 141: 220-233 (2007). Burns LR, Lee JA, Bradlow ET, Antonacci AC. "Assessment of Medical Devices: How to Conduct Comparative Technology Evaluations of Product Performance." International Journal of Technology Assessment in Health Care. 23(4): 455-463 (2007).