The effect of comparative effectiveness research on drug development innovation: a 360° value appraisal

John J Doyle
Quintiles, NY, USA

Abstract: The drug development process is in dire need of transformation. Even after achieving regulatory approval, pharmaceutical companies are increasingly seeing their products subjected to health technology assessments (HTAs) by public and private payers. The cornerstone of HTA value appraisal, and thus reimbursability, is comparative effectiveness research (CER), a ‘real-world’ comparison of a new product with the existing standard of care. Burgeoning demand for CER will fundamentally transform drug development by forcing biopharmaceutical manufacturers to view drug innovation from a holistic, 360° perspective. Specifically, drug and device developers must alter their existing approach to R&D by: adapting experimental research design methods to address multiple stakeholder demands; demonstrating real-world value through a suite of post-market observational research methods; and creating a transparent CER evaluation protocol based on standard principles. In the long-term, CER is forecast to propel innovation by focusing R&D on products that deliver real-world value to multiple customers and market stakeholders.

Keywords: comparative effectiveness research, comparative effectiveness balance sheet, health technology assessment, large simple trial, moderator, observational studies, patient randomized clinical trial, registry, ‘real-world’ value

Comparative effectiveness research

The drug development process is in dire need of transformation. As the cost, time, and complexity involved in developing new products continue to rise sharply – to as much as US$2 billion over 8 or more years – this transformation is necessary for both the pharmaceutical industry’s survival and to serve broad public health goals. Today, only one-quarter of novel drug candidates in Phase II advance successfully into Phase III. Attrition amplifies in late-stage development, with a further one-half of candidates failing to reach submission status. Even after achieving regulatory approval, developers are increasingly seeing their products subjected to health technology assessments (HTAs) by public and private payers.

The cornerstone of HTA value appraisal, and thus reimbursability, is a ‘real-world’ comparison of a new product with the existing standard of care, known as comparative effectiveness research (CER). This concept is defined by the Institute of Medicine as the ‘comparison of effective interventions among patients in typical patient care settings, with decisions tailored to individual patient needs’.1

CER weighs the benefits and harms of various modalities used to prevent, diagnose, treat, or monitor clinical conditions to determine which work best for particular types of patients and in different settings and circumstances. Conceivably, the refinement
of CER methodology and advancement of the science could harness the power of ‘crowdsourcing’ – defined by Howe, who coined the term, as ‘the application of open source principles to fields outside of software’ – by involving a broad range of researchers. Involving scientists in public, private, and academic domains will result in best practices in design, implementation, analysis, reporting, and communicating across those domains.

Results from CER studies can help patients, clinicians, policymakers, and purchasers make more informed decisions, ultimately improving care. Health experts and policymakers anticipate that CER will yield greater value from America’s healthcare system and better outcomes for patients. Despite spending more on care than any other industrialized nation – US$2.4 trillion in 2008 – the United States lags behind other countries on many measures of health, such as infant mortality and chronic disease burden. In a recently published article, Pearson and Bach posit that ‘the possible advantages of using comparative effectiveness research to help set reimbursements for newly covered services … is a promising option that we would argue the nation cannot afford to ignore’.

Successfully communicating the results of CER to all stakeholders will be essential if behaviors and practices are to be changed. The challenges of translating a CER finding to a change in practice is demonstrated by the Anti-hypertensive and Lipid-Lowering Treatment to Prevent Heart Attack Trial (ALLHAT), which found that diuretics were more effective than newer drugs in stemming heart problems. Despite efforts to communicate the results, this trial appears not to have had a great effect on diuretic use.

In 2003, Section 1013 of the Medicare Modernization Act ruled that the Agency for Healthcare Research and Quality (AHRQ) should evaluate the comparative clinical effectiveness of technologies. Additionally, the American Recovery and Reinvestment Act of 2009 (ARRA) provided US$2.4 billion in 2008 – the United States lags behind other countries on many measures of health, such as infant mortality and chronic disease burden. In a recently published article, Pearson and Bach posit that ‘the possible advantages of using comparative effectiveness research to help set reimbursements for newly covered services … is a promising option that we would argue the nation cannot afford to ignore’.

Successfully communicating the results of CER to all stakeholders will be essential if behaviors and practices are to be changed. The challenges of translating a CER finding to a change in practice is demonstrated by the Anti-hypertensive and Lipid-Lowering Treatment to Prevent Heart Attack Trial (ALLHAT), which found that diuretics were more effective than newer drugs in stemming heart problems. Despite efforts to communicate the results, this trial appears not to have had a great effect on diuretic use.

The Patient Protection and Affordable Care Act explicitly states that: ‘The Patient-Centered Outcomes Research Institute … shall not develop or employ a dollars per quality adjusted life year … as a threshold to establish what type of healthcare is cost effective or recommended’. However, a recent paper argues that ‘a ban on valuing life extension presents its own ethical dilemmas. Taken literally, it means that spending resources to extend by a month the life of a 100-year-old person who is in a vegetative state cannot be valued differently from spending resources to extend the life of a child by many healthy years’. In fact, cost has been used implicitly to prioritize which products are subject to comparative effectiveness research. For example, many observers believe that the comparison of ranibizumab and bevacizumab for macular degeneration was initiated because ranibizumab is priced many times higher than bevacizumab; the study itself did not examine cost.

Burgeoning demand for CER will fundamentally transform drug development by forcing biopharmaceutical manufacturers to view drug innovation from a holistic, 360° perspective. Specifically, drug and device developers must alter their existing approach to research and development in three ways:

1. **Adapt experimental research design methods** to address multiple stakeholder demands.
2. **Demonstrate real-world value** through a suite of post-market observational research methods.
3. **Create a transparent CER evaluation protocol** based on standard principles.

### Adapting experimental research design methods

Drug development must be transformed to circumscribe a product’s value proposition from multiple stakeholder viewpoints, and implement the optimal design to meet their needs (Figure 1). This 360° view of a product requires high-value innovation – products that improve therapeutic benefit, safety, quality of life or convenience – and not just new medicines. Complicating this 360° view further is the constant evolution of stakeholder perspectives, and the need to adapt the presentation of a therapy’s value proposition in response. Accordingly, value adaptation demands better trial design to provide appropriate data for each stakeholder group.

A useful approach will be the large simple trial (LST), which includes many (typically several thousand) people and extends over a long period of time. LSTs have broad eligibility criteria, simple enrollment procedures, collect minimal data, and use major illness or death as endpoints. They are
Effect of CER on drug development

intended to study treatments under ‘real world’ conditions of use.\textsuperscript{11} The study design can be viewed as a hybrid between a randomized clinical trial (RCT) and an observational cohort study, with randomization to avoid confounding, and endpoints evaluated mainly through observational follow-up. An example is the VOLUME Trial to evaluate the long-term pulmonary and cardiovascular safety of an inhaled insulin powder.\textsuperscript{12}

The advantages of an LST are inclusiveness, relevance to the usual practice setting, and naturalistic approach to therapy.\textsuperscript{13} LSTs have great promise in answering CER and safety questions, and can complement RCTs.

Historically, biopharmaceutical products were launched into the healthcare system based on cross-sectional appraisal by regulators. CER now demands a much more complex valuation scheme – one that is multi-dimensional and longitudinal. With real-time information exchange and ever-increasing connections within and between stakeholder groups, product evaluation will persist throughout its lifecycle. It is in the best interest of biopharmaceutical companies to pressure-test their product profiles early in development from many vantage points and simulate real-world performance using the best data available to guide evidence-based product design.

In the future, stakeholders will increasingly require customized clinical and commercial evidence derived from patient-level analysis of real-world outcomes data. This reflects the fact that experimental efficacy and safety translate into observed benefits and risks, once a product is launched. All these evolving factors translate into many varying vantage points on CER both between and within stakeholder groups.

Patients

This group’s perspective on ‘high value’ involves the traditional metrics of safety and efficacy, plus humanistic factors such as quality of life improvements and convenience of care. The required out-of-pocket costs under health insurance schemes may also be a factor. Personalized medicine will be a focus of interest, including information on which therapies work best for which specific patient populations. A recent paper indicates that at present, stakeholders have a poor understanding of evidence-based healthcare, highlighting a need for education.\textsuperscript{14}

Biopharmaceutical industry

CER is inextricably linked to progress in biomedical and biopharmaceutical research, helping to quantify the impact and value of innovation. As the use of CER is expanded, the cost of introducing innovations to market may increase in some areas, but decrease in others. Traditional RCTs will keep their place as the gold standard and cornerstone for evaluating safety and efficacy, but will be complemented by observational research to assess benefit, risk, tolerability, and value. It is thus critically important that RCTs be done well and done with an eye toward identifying safety as well as benefit ‘signals’ and against the background ‘noise’ in real-world healthcare systems.\textsuperscript{15}

When evaluating observational data, one needs to be cognizant of bias, confounding, and effect moderation. In addition to addressing these issues with study design levers, such as inclusion/exclusion criteria and matching, various statistical techniques should be considered, including multivariate modeling, propensity scoring and moderator analysis.

Physicians

CER has potential to help physicians make evidence-based recommendations to accommodate the unique characteristics and circumstances of each patient.\textsuperscript{16} Where multiple options exist, CER can also help physicians determine which interventions and strategies are most effective, safest or least costly. A recent analysis by the Medicare Payment Advisory Commission (MedPAC)\textsuperscript{17} found that physicians broadly agreed that CER is useful, noting the importance of subpopulations and variances in treatment response. Communication and education around CER findings must be improved; for example, there was little impact on practices after publication of the findings of the Cardiac Arrhythmia Suppression Trial (CAST),\textsuperscript{18} which failed to demonstrate a benefit for certain anti-arrhythmic drugs, and the Clinical Antipsychotic Trials Intervention Effectiveness (CATIE) trial,\textsuperscript{19} which questioned the effectiveness of newer antipsychotic medications.\textsuperscript{20}
Payers
Managed care organizations are reported to universally favor CER,23 which has potential to provide solid information to support value-based purchasing and benefit designs. Payers need CER to help guide decisions such as whether to elevate a treatment to first-line status, which formulary tier to assign it to, and what requirements to establish for prior authorization, step edits, and other utilization control techniques at their disposal. If evidence is lacking, other factors may dominate these coverage decisions, such as cost concerns or fears about potential negative publicity. If new therapies are priced at parity with existing treatments or shown to be of equal or better cost-effectiveness, payers might simply allow physicians and patients to sort out preferred therapies over time with the help of post-marketing observational data.

An example of an ongoing CER study that focuses on a specific sub-population of patients is the Genotype Guided Comparison of Clopidogrel and Prasugrel Outcomes (GeCCO) trial, a prospective observational cohort study.22,23 While GeCCO is designed to compare the effectiveness of the two drugs, the results could have far-reaching ramifications for patient safety and significant cost implications for health plans that pay for these drugs.

Some payers are explicit about their requirements; for example, in May 2010, WellPoint, Inc. became the first health benefits company to develop standardized CER guidelines that it will use in formulary decision making.24 The WellPoint Outcomes-Based Formulary will use clinical efficacy and clinical effectiveness ‘real world’ data to make formulary decisions with the help of ‘1) improve clinical health outcomes; 2) improve quality of life; 3) improve productivity at work, school, and leisure activities; and 4) reduce total cost of care (pharmacy and medical)’.25 The company’s guidelines explicitly note that ‘a more expensive medication can be less expensive overall if the member’s health is improved, resulting in use of fewer healthcare resources’.

Policymakers
Because federal health insurance programs play such a large role in financing medical care and account for such a large share of the federal budget, the government has a major interest in CER.26 In the future, this will be driven by the Patient-Centered Outcomes Research Institute (PCORI), which will prioritize and invest in CER using a largely stakeholder-driven process. This approach has potential to influence inefficient providers so that they adopt cost-effective practices instead of cost-ineffective ones.27 To address various stakeholder demands for optimal product performance, it is important to understand the various checks and balances that are in play. Figure 2 illustrates the evidence continuum for a biopharmaceutical. The double-helix in the center of the chart represents the dynamics of the healthcare system – a modulating process that influences how and where a particular product is used, and therefore how the product performs. While the pharmaceutical industry has historically approached development with a laser-focus on regulatory approval, the advent of CER means that a broader set of stakeholders are demanding diverse, real-world outcomes data and relative benefits, and biopharmaceutical manufacturers must incorporate these demands into the research and development process. Risk factors that moderate the effect of a product are important to consider in any of the three main methodologies for CER (observational studies, trials, and systematic reviews), not just in observational studies. Additionally, in observational research, there is a need to consider alternative explanations of effects such as confounding, bias and chance, as alternative explanations of the drug’s effect.

The level of sophistication of analysis increases from left to right in Figure 2, starting with a traditional safety profile, then tracking signals for adverse events, and then proactively designing studies to evaluate the product’s risk-benefit profile. For the value evidence channel (Figure 2A), it is important to examine the treatment efficiency and effectiveness (including patient-reported outcomes) in order compare the clinical, economic, and humanistic value to the standard of care. For the process evidence channel (Figure 2B), the treatment process needs to be evaluated from a healthcare systems perspective to understand interrelationships between stakeholders. Elements include care mapping, decision analytics and process optimization. For the safety evidence channel (Figure 2C) elements include pharmacovigilance, pharmacoepidemiology, and risk modeling. Risk modeling allows for profiling of patients, as well as the evaluation of their chance of receiving

![Figure 2 Evidence-driven value chain.](https://www.dovepress.com/)

Stakeholders in the new health landscape will require customized clinical and commercial evidence derived from patient-level analysis of real-world data.
treatment and the likelihood of adverse or beneficial outcomes. Moving from left to right, efficacy is ultimately translated into patient benefit, while safety is translated into risk; this overall real-world risk-benefit profile must then be communicated to all healthcare stakeholders.

CER has potential to inform decisions throughout the product life-cycle (Figure 3), and should be incorporated from early development onwards, through to patent expiry and beyond (when it can be instrumental in identifying new indications and subpopulations). Overall, the increasing focus on CER will trigger adjustments to research design including:

- Uptick in head-to-head comparisons
- Addition of composite endpoints, including cost-effectiveness
- Broadening of sub-group analysis to identify niches for optimal risk-benefit
- Longer trials to track downstream outcomes
- Increased sample size to accommodate additional endpoints and patient sub-groups.

**Demonstrating real-world value**

The various healthcare stakeholders – including physicians, patients, payers, providers, and policy-makers – each define value differently. Any assessment of real-world value must therefore allow the biopharma sector to actively monitor, measure, and report on the performance of its products in a naturalistic setting.

Demand for CER will require drug and device developers to provide further evidence of a product’s value in the post-marketing space through the conduct of various types of observational studies. Observational studies typically involve a much larger study size than randomized controlled trials and may run for many years, giving a truer picture of a drug’s performance under real-world conditions. CER may also trigger downstream ‘check-points’ for companies to revisit the risk-benefit (and related cost-benefit) profiles of their drugs, thus supporting a 360° perspective on value appraisal.

This risk-benefit evaluation deserves increased attention, since successful delivery of an optimized risk-benefit balance typically links directly to improved cost-benefit performance.

There is a need to evaluate a drug’s performance in various sub-populations that may carry risk factors – such as age, other medicines being taken or comorbidities – moderating the effect of the product. It is important to understand how those moderators cluster in the population, to allow targeting of populations where the benefit-risk of the product will be optimized.

Unlike randomized trials, which can be analyzed using a range of standardized, critical appraisal tools, there are at present no universal guidelines for evaluating the quality of observational research. One effort to address this is the STROBE initiative, Strengthening the Reporting of Observational Studies in Epidemiology, which offers a 22-point checklist of factors to include in an accurate and complete report of an observational study. Designing, implementing, analyzing, and reporting observational studies in the absence of such guidelines will undermine the very objective of the research of gaining clarity on biopharmaceutical performance post-launch.

Another initiative to evaluate risks and benefits is the Observational Medical Outcomes Partnership (OMOP), a public-private methods development and testing consortium, which is taking a two-tiered approach involving the OMOP Research Core and the OMOP Extended Consortium. This aims to serve public health by testing whether multi-source observational data can improve our ability to assess drug safety and benefits.

There is considerable overlap between observational research and pharmacoepidemiology, the scientific backbone of therapeutic risk management. In an effort to help investigators plan, conduct, and evaluate pharmacoepidemiologic research, the International Society of Pharmacoepidemiology (ISPE) published Guidelines for Good Pharmacoepidemiology Practices. These address protocol development, responsibilities, study conduct, communication, adverse event reporting, and archiving. An earlier United States Food and Drug Administration (FDA) document, Guidance for Industry: Good Pharmacovigilance Practices and Pharmacoepidemiologic Assessment, had similar objectives, and also included guidance on safety signal identification, pharmacoepidemiologic assessment, safety signal interpretation, and pharmacovigilance plan development.
Patient registries are prospective, observational cohort studies of patients with, or at risk for, a particular disease and/or receiving a particular treatment/intervention. They can be used for understanding natural history, assessing or monitoring real-world safety and effectiveness, assessing quality of care and provider performance, and assessing cost-effectiveness.\textsuperscript{33} Registries can help in evaluating safety signals, part of the safety evidence channel in the evidence-driven value chain (see Figure 2C). These signals are identified from spontaneous case reports, published literature or other sources, and identifying factors affecting the risk of adverse outcomes, such as dose, timing of exposure, or patient characteristics. Surveys of patients or healthcare providers can gather information on topics such as safety signals, knowledge about labeled adverse events, and use of a product as labeled. The FDA Sentinel Initiative, for example, is focused on real-world long-term safety and risk data based on retrospective analysis of claims data.

In parallel, the Centers for Medicare and Medicaid Services (CMS) is focusing increasingly on evidence of value — as manifested by its national coverage decisions that recommend ‘coverage with evidence development’. Each of these two approaches, with a focus on safety and risk by the FDA, and value by CMS, examines only half of the picture. A provocative approach would be to link the two by creating a common channel to measure the multiple dimensions of a product’s real-world performance. This would allow feedback between the safety and value analyses, helping to elucidate the risk-benefit, and cost-benefit, of a product using the best observational science.

**Creating a transparent CER evaluation system**

Before CER can be widely adopted by the industry — allowing drug development to transform in response to stakeholders’ 360° view — there is a market need for a transparent evaluation protocol based on standard principles. As WellPoint Chief Pharmacy Officer Brian Sweet said in May 2010: ‘Currently, CER evaluation is much like the wild, wild West. While a few organizations have made some strides, there is little in the way of publicly available, comprehensive evaluation criteria’.

This needs to change. CER should focus on a holistic notion of evidence development, encompassing a variety of treatment process and outcome metrics, measured by alternative research designs. The comparison of competing therapies should be based on similar types of real-world data. Indeed, comparing the limited information on the package insert of one product with real-world experiences with another would not give a fair evaluation.

A standardized method is needed to accurately and fairly appraise the value of pharmaceutical innovation from multiple stakeholder perspectives, and to facilitate comparison of competing agents. Financial valuation techniques may be referenced as a conceptual framework for biopharmaceutical valuation. Financial analysts use a variety of measurement techniques to evaluate products, services, and companies. These analysts measure company value under conditions of uncertainty to provide insight on investment opportunities. Similarly, CER analysts aim to measure product performance under conditions of uncertainty in order to provide information on the public health impact of adoption of biopharmaceuticals. This measurement should include both short- and long-term evaluations of benefits and risks.

A ‘comparative effectiveness balance sheet’ for biopharmaceutical products can be conceptualized as a snapshot of positive and negative product attributes that is updated periodically. At launch, the ‘actual’ data driving the CER balance sheet will be based on evidence derived from the randomized controlled trials conducted in clinical development. For example, any patient populations not studied adequately in the RCT setting yet included explicitly or implicitly in the label should be noted as ‘pro forma’ performance estimates. Only when those patient subpopulations are studied comprehensively should they be labeled as ‘actual performance’.

For a standardized, universal system to be developed, the overall constellation of product attributes would need to be weighed. Before attempting to assign a value to a product, it is important to understand what really matters to key stakeholders — particularly payers, given their influence in the system. When making decisions about therapies for a particular indication, what really matters to them? Is it morbidity or mortality concerns? Is it unmet medical need? Is it patient convenience? Each of these factors requires a relative weight, so it is important to understand what matters most for that particular indication. The idea is to get a scorecard of what matters to payers and other stakeholders at that particular time, and to use this as a framework for product valuation and construction of the comparative effectiveness balance sheet. One measure that could be used as a common denominator for risks and benefit valuation is the quality-adjusted life year. Clearly, more research is needed to investigate potential metrics for standard measures of product effectiveness in order to facilitate fair comparison.
Conclusion
In the future, CER may begin to gain traction on a larger scale as the drug registration process is accelerated, resulting in payers requiring additional evidence of efficacy before fully covering a drug. In the long-term, CER should propel innovation by focusing R&D on products that deliver real-world value to multiple customers and market stakeholders.

Biopharmaceutical innovation can only translate to improved public health through a coordinated effort. This needs to involve all stakeholders and focus on delivering the optimally balanced risk-benefit treatment to the most appropriate array of patient populations identified in our rapidly changing healthcare system.

There would be a clear benefit to all stakeholders – and to public health – of linking the FDA’s safety/risk-focused initiatives with the CMS’ investigations of value through CER, so that the two avenues of research could converge to inform one another. Companies should capitalize on this wider market interest in their products’ worth, surround their technologies with formidable evidence of value, and be properly rewarded for their innovations.

Disclosure
The author reports no conflicts of interest.

References
8. The Patient Protection and Affordable Care Act. PL 111-148. 3-23-2010.