

Comparative Effectiveness Research: International Experiences and Implications for the United States

by Kalipso Chalkidou, M.D., Ph.D., and Gerard Anderson, Ph.D.

July 2009

Introduction

Overview

The idea of developing a comparative effectiveness (CE) program in the United States is not new. More than 15 years ago, the Office of Technology Assessment (OTA) published a report discussing the rationale and obstacles associated with the creation of a comparative effectiveness entity. “The justification for most medical practices used in the United States today rests on the experience and expertise of clinicians and patients rather than on objective evidence that these practices can measurably improve people’s health. Compiling objective evidence is considered by some...highly controversial, because the evidence might be applied in ways that would limit individuals’ choices of medical treatments.”¹

To date, comparative effectiveness research activity in the United States has been limited in its funding and scope, although that appears to be changing given its prominence in the American Recovery and Reinvestment Act. Other countries, who have been using comparative effectiveness to make decisions for some time, may offer instructive examples for policymakers in the United States.

In this report, we examine how CE programs operate in other countries. We focus on four policy issues and how various countries have addressed them. We reviewed CE programs in several industrialized countries (primarily the United Kingdom and Germany) but also included the experiences of several middle-income countries. While CE research must recognize country-by-country differences in medical practices and ability to pay, much of the clinical information is based on international peer-reviewed literature, thus making CE information applicable internationally. As a result, several middle-income countries have been able to take advantage of the available information to introduce

innovative CE programs. We discuss programs operating in:

- Australia
- Brazil
- Canada
- Germany
- England and Wales
- France
- Russian Federation
- South Korea
- Sweden
- Turkey

Focus of Report

Consistent with the CBO definition of CE (Appendix 1) and with international experience, this report focuses on medical technologies, including drugs and devices, surgical interventions, and diagnostic tests. Some CE entities focus only on drugs (Australia) while others focus solely on medical devices and procedures (Ontario). The UK also conducts CE for public health interventions and programs. The present analysis, however, does not consider CE’s role in public health. Although some have proposed that CE entities should also address the CE of alternative service delivery models, these models are beyond the scope of this analysis and are generally not part of the domain of CE programs in other countries.

This report focuses on four main topics:

- Establishment and evolution of comparative effectiveness agencies
 - What was the original reason for the establishment of the CE organization?
 - How is it organized and related to other entities?
- Governance and accountability arrangements
 - What is the CE organization’s legal status?

- How is it governed?
- How is it funded?
- What is its size and cost?
- How are its decisions appealed?
- Methods to inform decision making
 - Who decides what topics to address?
 - Who conducts analyses?
 - What are the primary data sources/ techniques?
 - How is the input of various constituencies included in the analysis or its interpretation?
 - Are costs an important consideration?
- Dissemination and impact
 - How is information disseminated?
 - Who enforces the recommendations, if at all?
 - Has the information led to changed medical practices or altered costs?

Core Attributes of Comparative Effectiveness Entities: Lessons for the United States

Before describing various core attributes of the CE entities studied here, we draw on international comparisons to propose some possible lessons for the United States.

Organizational Forms

CE entities were either established as part of a comprehensive policy strategy or rapidly integrated into the overall health system. What differentiates CE from conventional medical research is that it is demand-driven rather than investigator-driven. Demand is defined by the needs of decision makers, including public and private payers, clinicians, patients, and professionals. Most CE entities were established as part of their respective

nations' comprehensive restructuring of their health care system. International experience demonstrates that the more "hard-wired" a CE agency into its nation's health care system, the greater is its impact.

Independence from government is important. While most of the CE entities examined for this analysis were the creation of their respective government and receive different levels of government funding, they all exhibit some level of independence from government.

Independent oversight boards are common. One or more independent boards of directors ensure good governance and accountability.

Some CE entities set standards while others only identify and recommend standards. Some of the CE programs only identify standards, leaving final decisions to a separate entity. More commonly, however, the CE entity sets standards under the supervision of an independent board of directors, with oversight/monitoring provided by another component of the health system.

Most CE entities are relatively small and operate with a modest budget. Most CE entities operate with a few hundred people and a budget under \$60 million. They outsource most of their research and rely on other entities to conduct analyses and disseminate results.

Methods

The prioritization process for selecting topics for review is critical. A prioritization process closely linked to the needs of decision makers guides the selection of topics for analysis. Increasingly, decision makers are interested in evaluating established practices with a view to eliminating wasteful activities.

Methodological rigor is critical to the acceptance of the CE entity. Methodological rigor forces the debate to focus on issues rather than on politics and helps protect the CE entity from the influence of vested interests. Buy-in from academic groups often promotes rigor.

Transparency is crucial. At the United Kingdom's National Institute for Clinical Excellence (NICE), all information that drives decisions, including CE evidence and committee deliberations, is available to the public. NICE scrutinizes all confidential academic and commercial submissions and encourages sources to release the information to the public as soon as possible. Other countries have adopted similar transparency principles.

Timeliness of decisions is important. Resource allocation decisions need to be made quickly. CE entities use data from a variety of sources and often must make tentative decisions until more information becomes available.

CE entities use different types of data to reach recommendations. Analytic approaches to making recommendations vary from primarily relying on randomized clinical trials (RCTs) to relying on claims data, meta-analyses, and other types of data. In addition, CE entities sometimes make use of manufacturers' data and unpublished data and sometimes re-examine existing evidence as new information becomes available.

There is increasing use of conditional coverage as a means of linking prospective evidence generation to decision making. CE entities engage in evidence synthesis as their main type of research largely because of resource constraints and the timeliness of current evidence. Increasingly, however, CE is being linked to prospective trials.

Applicability to different populations is essential. International experience shows that the application of methodological tools such as individual patient data pooling from trials or decision analyses based on observational and RCT data ensure that studies are relevant to given subpopulations. In the UK, the need for subpopulation-specific information has triggered public investment in methodological research and human capacity building through formal training programs in UK universities.

Costs are an important consideration. Some CE programs did not initially consider costs, although all do so today. Countries have found different ways to incorporate costs. England's approach is the most explicit, specifying a cost-effectiveness cut-off range (threshold range). Australia is less clear about having an explicit threshold, but open about costs being a key consideration in decision-making. In fact, Australia was the first country to include such cost considerations. Ontario is undergoing a gradual move toward consistent consideration of costs. Germany is in the process of devising a completely new method for economic evaluation that differs from standard cost-effectiveness analysis and allows each sickness fund (insurer) to set its own threshold.

Economic evaluations are often included in comparative clinical trials even in settings where cost-effectiveness is not a requirement. While some stakeholders may resist the incorporation of cost-effectiveness considerations into the decision-making process, countries have yet to tackle resistance to cost considerations through the introduction of some type of "firewall" that isolates clinical from cost-effectiveness issues. Instead, CE entities have gradually assumed responsibility for economic analysis as part of a longer-term maturation process. Despite the many perspectives for measuring costs (societal, patient, insurer, and government), most CE entities focus on the payer perspective.

Dissemination and Impact

Inclusiveness of all stakeholders, including manufacturers, is critical. Gaining stakeholder buy-in is essential in ensuring that CE-based decisions have the desired impact. Inclusiveness depends on engaging all stakeholders not only in the decision-making process but also in methodology design.

Most CE entities are not charged with saving costs. Most CE entities were not initially charged with cost containment as their primary goal. In fact, most CE programs have been responsible for health care cost increases in that they approve new technologies that are more expensive than the technologies being replace. Because CE programs are part of a larger health system, it is difficult to assess what cost savings have been realized or to estimate what overall spending might be if CE entities were not in place. Furthermore, most agencies have been operating for only a short period, making it difficult to assess their impact on longer-term spending and health outcomes. Nonetheless, general speculation holds that CE entities are responsible for allowing payers, consumers, and industry to develop a better understanding of health outcomes, increase the transparency of health care investment, and reduce regional variation.

An appeal process is essential. Many countries have established an appeals mechanism for decisions made by the CE entity; the format varies from country to country.

Strong political endorsement can be critical. The CE entity often makes decisions that are unpopular with some constituencies. In all the international examples studied here, the country's central or provincial government grants full support to the CE entity in nearly all cases. The political and legislative support received by CE entities means that they can withstand the political and lobbying

pressures of dissatisfied stakeholders, including manufacturers, patients, or professionals. In isolated instances, the government has tried to interfere or influence a CE entity's decisions, but such cases have generally occurred only once.²

Most CE entities rely on other entities to enforce and disseminate their findings. Most CE entities cannot force stakeholders to implement their recommendations and instead rely on other entities to do so. Dissemination and implementation support are increasingly important components of CE, and CE entities are gradually assuming responsibility for such components.

Why Have Countries Created Comparative Effectiveness Entities?

In most cases, the creation of a CE entity represented a response to concerns about obtaining value, limiting unwarranted variation in medical practice, and the cost and diffusion of new technologies. Cost containment was not part of the original impetus for CE entities. In most countries, the CE program was generally part of a larger reform effort.

England (UK)

The British National Health Service (NHS) was established in 1948 as a publicly financed, single-payer system providing free universal access to health care. Following a period of perceived chronic under funding of health care relative to other industrialized countries, the NHS benefited from the Labour government's (under Tony Blair) commitment to increased public funding as of 2000. Between 2002 and 2008, the NHS experienced a planned 50 percent increase in real spending that was designed to bring health spending levels closer to the European Union average.

During this period, the government introduced a series of programs that modified operation of the NHS. As described in a white paper announcing the restructuring of the NHS, the National Institute for Clinical Excellence (NICE) was part of a comprehensive strategy aimed at better informing practice and policy decisions.^{3,4} According to the white paper, "The Government is determined that the services and treatment that patients receive across the NHS should be based on the best evidence of what does and does not work and what provides best value for money. . . . At present there are unjustifiable variations in the application of evidence on clinical and cost-effectiveness. . . . The Government will spread best practice. . . . by establishing a new National Institute of Clinical Excellence which will promote clinical and cost-effectiveness by producing clinical guidelines and audits, for dissemination throughout the NHS."⁵

In the same white paper establishing NICE, the government reinforced its support of the NHS research and development program to undertake research assessing the clinical and cost-effectiveness of health technologies. The white paper also announced the launch of a new "horizon scanning" function "for emerging clinical innovations. . . to help set research priorities, to provide information for planning services, and to identify the need for clinical and service guidelines which the new National Institute may be commissioned to develop." These new government initiatives were expected to work collaboratively to rationalize health allocation decisions.

Germany

The German health care system dates to the end of 19th century amid the Bismarckian reforms that created social health insurance. The system is decentralized, with more than 90 percent of the population covered through mandatory social health insurance offered

by several sickness funds. High-income individuals may opt out of the system and purchase private health insurance. The presence of several payers may have influenced the creation, organization, and scope of Germany's CE entity.

In the early 2000s, Germany set forth several objectives in reforming its health care system: consolidating insurance funds and broadening their financial base; reducing fragmentation in financing streams and increasing the proportion of tax funding; imposing a single premium (as a proportion of income); and offering citizens the right to choose their insurance fund and, through choice, introducing regulated competition between funds.⁶ The 2004 creation of the German Institute for Quality and Efficiency in Healthcare (IQWiG) was part of the overall reform strategy. Initiating operations in 2005, IQWiG introduced cost-effectiveness and clinical effectiveness as conditions for evaluating coverage and reimbursement of health technologies.

At the outset, IQWiG's main role was to identify national quality standards by collating and presenting to decision-making bodies an assessment of the evidence of the comparative clinical effectiveness of various medical interventions. German policymakers saw a pressing need for an independent "standard-identifying" (rather than "standard-setting") body to inform negotiations between insurers and professionals. IQWiG therefore transmits its information to another decision-making body that comprises sickness funds and providers. This approach stands in contrast to the UK's approach, whereby some types of NICE guidance are mandatory (discussed later in more detail). In the 2007 reforms, IQWiG's responsibilities were expanded to incorporate cost-effectiveness data into its assessment portfolio, although the means of such incorporation is still under discussion. IQWiG is not expected to issue cost-effectiveness advice before 2009.

Australia

Australia's Pharmaceutical Benefits Scheme (PBS) has its antecedents in a limited program established after World War I to provide pharmaceuticals free to returning veterans. In the following decades, the idea of a system to provide medications to all Australians gained increasing acceptance and was adopted during World War II. In 1953, the government granted the Pharmaceutical Benefits Advisory Commission (PBAC) authority to recommend drugs for addition to the formulary.⁷ Initially, PBAC based its coverage recommendations for drugs solely on perceptions of clinical need. "It is not the intention to limit the inclusion of any drug or compound on the score of cost. The only question asked will be: Does it contribute anything towards the medical efficiency of the formulary?"⁸

Operating with only limited transparency, PBAC initially consisted of senior public servants directly appointed by the Minister for Health. Gradually, however, as cost pressures and patient expectations grew, PBAC changed its administrative style. First, operation of the committee became more transparent, and membership expanded to draw from the clinical and academic communities rather than only from the government civil service. In the 1970s, the committee started making the reasons for its decisions known to the medical community and then to the broader public. Second, as cost became a concern in the 1960s and 1970s, PBAC introduced a means-tested patient co-payment. In the late 1980s, PBAC became the world's first reimbursement committee to consider drug costs directly in its decision making. As of 1993, PBAC considers comparative value in its decision making in accordance with the principle of "purchasing outcomes"; that is, a higher price is acceptable only when a drug offers additional clinical benefit (either greater efficacy or reduced toxicity or both) compared to current alternatives.⁹

Given that Australia operates a positive listing system, PBAC must consider all new drugs, and its approval is a prerequisite for listing a drug on the national formulary. It does not make pricing recommendations but instead considers the comparative clinical and cost-effectiveness of technologies based on the unit price proposed by the manufacturer and then makes recommendations to the pricing authority. The authority in turn advises the government as to the most appropriate price.

France

The French High Health Authority (HAS) was established in 2005 as an independent public body charged with a wide range of responsibilities, including provider accreditation, guideline development, definition of the basic package for chronic disease patients, promotion of information technology tools in health care, and informing the French insurance system's decisions about listing and reimbursement for medical technologies and services. Although not explicitly stated in its responsibilities, one of HAS's political objectives was to help rationalize spending on health care through quality improvement¹⁰ and, since January 2008, through adoption of cost-effective technologies.

Middle-Income Country Examples

Several middle-income countries are creating CE entities. As a starting point, they draw from the guidelines created by NICE, IQWiG, and other CE programs and then adapt them to their own setting, subject to their ability to finance interventions and management practices. They illustrate additional reasons for establishing CE entities.

Turkey is undertaking a series of significant reforms of its health care system. In particular, it is establishing a national entity responsible for reviewing the evidence of what works in health care in order to set quality standards, provide professional guidance, and inform coverage decisions as part of its guaranteed package of services

(including drugs). Turkey's political commitment to improving quality while guaranteeing a package of services seems to be driving the government's interest in using CE evidence to guide day-to-day coverage decisions.

In many countries, an important impetus for CE is guidance in the design of benefit packages. Increasingly, countries rely on CE to ensure that the services promised in the benefit package are accessible. Many Latin American countries, for example, face the challenge of a particularly generous state-promised benefit package in a climate of limited financial resources such that few drugs are available when a patient arrives at an outpatient clinic or hospital. Countries adopt CE to ensure that listed drugs are in fact available.

Brazil, for example, uses the NICE guidelines as the first step in a process to ensure that services in the benefit package will be available. Brazil's Ministry of Science reviews the NICE guidelines in the context of the country's medical care system while Brazilian health economists evaluate the cost of the NICE recommendations in the Brazilian context. This information, in combination with the requests of physicians, hospital managers, and patient advocates, is presented to the Minister of Health for a final decision on what should be included in the final benefit package. Once a drug or device is included in the benefit package, the expectation is that it will be available to all patients.¹¹

In the case of the Russian Federation, different parts of Russia (oblasts) have used CE information to make resource allocation decisions and improve technical efficiency. The current system of resource allocation is highly centralized and heavily hospital-oriented. CE offers a means to improve allocative efficiency by reducing hospitalization, expanding the outpatient drug benefit, and setting professional standards for best practice.

Governance and Accountability Arrangements

CE entities' governance and accountability arrangements vary with respect to legal status and statutory functions; funding mechanisms; the size and cost of operations; collaboration with other entities in gathering and disseminating information; and appeal processes.

England (UK)

Legal status, funding, and statutory functions. The UK's NICE is a Special Health Authority that operates under a 12-member Board of Directors that consists of eight non-executive directors and a chair, all appointed by the independent Appointments Commission, and four executive directors, all appointed by the non-executive directors. The board determines policy and strategy and holds the executives to account for the achievement of objectives specified in the strategic and business plans. Among several subcommittees are audit, risk management, and citizens' council subcommittees. NICE also has a Partners' Council of major stakeholders appointed by the secretary of state. Stakeholders include representatives of patient groups, the health professions, NHS management, industry, and trade unions. The Partner's Council meets annually to review NICE's annual report and provide strategic advice to the institute.

NICE is funded by the Department of Health and issues its guidance directly to the NHS Health Ministers responsible for its work program (topic selection), although the institute is administratively independent of the Department of Health. Government departments and other stakeholders may participate in the guidance development process through designated consultation periods. However, reliance on government funding has led some to question NICE's independence from political pressures within NHS, especially in the early years of its operation.

Size and budget. When NICE was first established in 1999, it employed six people, including the directors of its programs, and operated with a budget of \$15 million. Over the years, the organization has grown and, in 2007, employed a staff of 270 with an annual budget of \$70 million. The Department of Health channels additional funding (approximately \$8 million) directly to the academic networks that undertake CE research to inform NICE guidance. NICE has always been a "virtual" organization and, despite its growth, still relies largely on a network of external experts, including members of the public, health care professionals, academics, and commercial entities, to identify, develop, and disseminate its recommendations. More than 2,000 people work with NICE over any one year to help develop guidance. Most volunteer their time and expertise.

Status of NICE guidance. The status of NICE guidance varies across the member nations of the UK, although the present analysis focuses exclusively on England, where NICE recommendations for the use of technologies became mandatory in July 2003.¹² Local public purchasers of care (Primary Care and Hospital Trusts) have a three-month period to identify the additional funding needed to implement NICE-recommended technologies. Given that NICE has no budget to fund its recommendations, Primary Care Trusts must identify additional funding in order to comply with the institute's recommendations. Increasingly, where NICE guidance applies, additional funding is made available for specific Health Resource Groups (the NHS equivalent of DRGs) in the form of "tariff uplifts." However, some argue that mandated NICE recommendations distort local priorities and crowd out non-drug, service delivery-type interventions for which such mandates usually do not apply. In response, NICE provides Primary Care Trusts with planning tools to prepare them for additional costs, and it increasingly focuses on identifying inefficient practice. In June 2008, an "NHS Constitution" that describes the basic standards of care to be expected

from the NHS was proposed. Access to NICE-approved treatments from its technology appraisal program will become a right if this legislation is enacted.¹³

To promote high-quality care by ensuring NHS compliance with national standards, including NICE guidance, the government established a monitoring body, known as the Healthcare Commission (HCC), at the same time that it established NICE. Currently, HCC reviews the performance of NHS providers against several core and developmental standards. The former are mandatory and include NICE guidance on the safety and efficacy of new interventional procedures and on NICE recommendations for specific health technologies. For the latter, NHS providers must be able to demonstrate implementation plans that include NICE clinical guidelines. In 2010, the HCC will be replaced by the Care Quality Commission, which will introduce a new system of registration standards for providers, including compliance with NICE technology appraisals and interventional procedure reviews.

Overall, instead of relying on rigid rules and regulations that force clinicians and public payers to adapt its recommendations, NICE relies on frontline NHS staff and patients for implementation of its guidance. It strives to gain stakeholder buy-in through the inclusive nature of its guidance development processes (see below for the Royal Colleges' participation in the process).

Data collection. NICE is a user of CE research to inform coverage decisions and to set quality standards across the NHS; it does not directly engage in comparative effectiveness research. It is heavily dependent on the UK's pre-existing infrastructure to produce needed evidence, mostly in the form of systematic reviews and evidence syntheses, in order to make coverage and policy decisions. NICE's data needs have resulted in increased

government funding of comparative effectiveness research through public agencies, namely, the National Institute for Health Research (NIHR).

The data collection program was established¹⁴ with the dual objectives to “commission research focused on the needs of patients and the public” and develop “evidence to inform and underpin health and social care policy.” NIHR includes a health technology assessment program and a horizon-scanning service, both commissioned by the Department of Health, to assist NICE with topic prioritization and evidence syntheses.

Further, NIHR has provided financial support to the Cochrane Collaboration, supports research into methods of critical appraisal and economic evaluation, and builds professional capacity through fellowship training programs across major UK universities. A recent reform of the public funding arrangements for supporting health-related methodological research has transferred responsibility for funding research on methodological issues to the Medical Research Council (MRC). The newly formed (2007) Methodology Research Program, with a total budget of \$120 million, aims to “support the development of new and improved systems and theories for health research” and is charged with providing support to decision makers such as NICE by addressing important methodological challenges.¹⁵ NICE now works closely with MRC to develop the research briefs that call for research into specific areas of interest to the institute, such as calibration of its cost-effectiveness threshold by using empirical evidence and the development of appropriate criteria for making conditional coverage decisions for new emerging technologies with a weak evidence base.

In addition to NHS research and development, the NICE clinical guidelines program is built on a pre-existing professional guidelines' development

network that encompasses England's Royal Colleges (equivalent to professional societies such as the American College of Radiologists). NICE maintains the network by contracting with the colleges and their affiliated teams of systematic reviewers and, more recently, with health economists to produce NICE guidelines, the full version of which is owned by the respective college. Each college maintains the right to distribute the full guideline that carries that college's logo. NICE may also use the material in the full guideline for its own purposes such as the development of implementation support or audit tools. The intellectual property arrangement between NICE and the Royal Colleges has expedited the establishment of the NICE clinical guidelines program—predominantly through the analytic work of the colleges, which agreed to buy into the NICE standardization and quality assurance processes—and has contributed to increased acceptance and implementation of NICE recommendations among clinical professionals who see the final output as a product of their own professional association rather than an externally imposed direction with little clinical input.

Appeal process. Stakeholders may appeal NICE recommendations for the use of technologies on the following grounds: (1) perversity in that no reasonable group of people would have formulated recommendations as presented by NICE, (2) violation of NICE procedural rules, and (3) violation of NICE scope of responsibilities. Appeals are heard by a panel consisting of at least two NICE non-executive directors and three other members with patient advocacy, industry, and NHS experience. Approximately one in three appraisals is appealed, with almost half of appeals upheld. If unsuccessful at the appeal stage, stakeholders may seek a judicial review of the guidance as has occurred only once thus far (Alzheimer's drugs); a decision is pending.

Germany

Legal status, funding, and statutory functions. To ensure stakeholder buy-in, IQWiG has assumed a strictly advisory role with respect to the Joint Federal Committee (JFC), which has final decision-making authority. JFC consists of providers (including hospitals and health care professionals) and insurance funds, with both groups equally represented. Patient representatives with no right to vote also sit on the JFC as does a “neutral vote,” which determines the final decision in the case of a disagreement between professionals/providers and insurers.¹⁶

IQWiG is an independent nonprofit non-governmental entity with a 30-member Board of Trustees, a Scientific Advisory Board, and a Steering Committee that includes the institute’s management and department heads and a Methods Group that is responsible for the methodology applied by IQWiG in its assessments and for institute processes.

The Federal Ministry of Health approved IQWiG’s establishment in July 2004,¹⁷ with the following responsibilities:

- Research, presentation, and evaluation of the current medical evidence base for diagnostic and therapeutic procedures for selected diseases;
- Production of scientific reports, expert opinions, and statements on the quality and efficiency of services delivered within the framework of the statutory health insurance system, taking into consideration factors concerning age, gender, and life circumstances;
- Assessment of evidence-based guidelines for the epidemiologically most important illnesses;
- Submission of recommendations on disease management programs;

- Assessment of the benefit and cost of drugs; and
- Provision to citizens of easily understandable general information on the quality and efficiency of health care services and on the diagnosis and therapy of diseases of substantial epidemiological relevance.

In 2007, the Ministry of Health added to the institute’s charge responsibility for making recommendations to insurance funds for the ceiling (maximum) price of new (or existing) listed technologies, based on evidence of comparative value. IQWiG has yet to issue any such recommendations and instead is developing its methodology for undertaking the required cost-effectiveness analyses.

Size and budget. IQWiG operates with a staff of approximately 90 and a budget of about \$30 million (2008). It is funded to “50% by a levy on every hospital case to be invoiced and to 50% by a tax on ambulatory and outpatient services reimbursed by statutory health insurance.”

Status of IQWiG guidance. JFC is not required to accept IQWiG’s advice but must explain its rationale if it decides not to do so. In a sense, JFC performs the “appraisal” part of the decision-making process, partly based on evidence collated and interpreted by IQWiG. IQWiG fits into a highly decentralized environment, with strong professional, union, and industry interest groups operating within a relatively opaque negotiation framework.

Data collection. IQWiG carries out CE in the form of evidence syntheses rather than through prospective studies, both in-house and through external groups (“collectives”) in Germany and other EU countries. External commissioning takes place mostly in disease areas in which IQWiG lacks relevant expertise. In such cases, IQWiG commissions external experts, mostly clinical specialists with critical appraisal skills, to collate

the relevant evidence by using IQWiG-designed “evidence table templates.” IQWiG undertakes (or commissions) little primary research in the form of prospective clinical trials and registries, although JFC may commission prospective studies in instances of important evidence gaps. Similarly, IQWiG does not perform claims data analyses, as is the case in other countries. Furthermore, given its primary reliance on meta-analyses of RCTs, IQWiG operates with a relatively underdeveloped network of external CE support structures compared to other similar entities abroad.

Appeals process. Relevant stakeholders may appeal IQWiG decisions.

Some Unique Aspects of CE Governance in Other Countries

Canadian Common Drugs Review (CDR) and Ontario Health Technology Advisory Committee (OHTAC)

In Canada, the provinces are responsible for managing, organizing, and delivering health care and for funding more than 80 percent of health care costs (public and private); the federal government funds the balance. Through the Common Drugs Review (CDR), which forms part of the Canadian Agency for Drugs and Technologies (CADTH), the federal government assesses drugs (with the exception of Quebec). In light of a CDR recommendation, each drug plan in Canada makes its own decision. Overall compliance with CDR guidance exceeds 90 percent; however, CDR decisions apply only to the Canadian provinces’ public programs, thus accounting for approximately half of overall spending on pharmaceuticals. To make its decisions, CDR relies on evidence submitted by the drug and device companies.

Responsibility for generating CE to inform coverage decisions for non-drug technologies tends to rest with the provinces (e.g., Institute for Health Economics in Alberta; Agence d’évaluation des technologies et des modes d’intervention en santé [AETMIS] in Quebec; and Medical Advisory Secretariat/

Ontario Health Technology Advisory Committee [MAS/OHTAC]). Ontario has developed a unique scheme whereby CE evidence, ranging from independent evidence syntheses to prospective trials and registries, is directly linked to coverage decisions. OHTAC, established in 2003 following an initiative led by Ontario hospital CEOs,¹⁸ uses dedicated networks of researchers, mostly funded by the provincial government, to evaluate health technologies upon request from potential technology purchasers, including the Ministry of Ontario or hospitals located in the province.

In the case of insufficient clinical and cost data to make fully informed decisions, Ontario has implemented a highly innovative and relatively unique model for generating and assessing evidence. Through what it calls conditionally funded field evaluations (CoFFE), OHTAC funds prospective trials that compare the clinical and cost-effectiveness of non-drug technologies.

Australian Pharmaceutical Benefits Advisory Committee (PBAC)

CE entities in Australia are part of federal or provincial governments rather than independent public agencies. Given that it operates a positive listing system, PBAC must consider all drugs licensed in Australia. Drugs rejected by PBAC are not listed on the formulary.

Methods to Inform Decision Making

CE entities must select technologies for review, appraise the technologies, and nearly always consider costs before making recommendations.

England (UK)

Topic selection. NICE plays an increasingly central role in selecting the technologies and medical conditions for which it issues guidance. As of 2006, NICE is responsible for identifying and selecting—according

to predetermined criteria¹⁹—high-priority topics and the most appropriate process (clinical guidelines, public health guidance, or technology appraisal) for developing guidance on these topics. The Minister of Health retains final responsibility for topics referred to NICE.²⁰

NICE relies on several sources to prioritize the selection of technologies and medical conditions, including a dedicated horizon-scanning service funded by the Department of Health and suggestions by the public through an open Web-based process. NICE places the recommended topics before panels led by the National Clinical Directors (government advisors) in the respective clinical areas (e.g., cardiovascular disease, public health, mental health, cancer). The panels draw their membership from professional associations, industry, academia, and the general public.

Calls by various stakeholders (including patient groups and industry) to allow more time to consider evidence, engage in the process, and appeal unfavorable decisions have led to a protracted selection process. Responding to concerns over the lengthening process, NICE has abbreviated some consultation rounds and accelerated the evidence analysis stage. While transparency and inclusiveness are important characteristics of the NICE topic selection process, the institute must balance its consultative responsibilities against the pressure to issue timely guidance.

Assessment and appraisal. NICE's charge is to help the NHS make evidence-informed decisions regarding the most clinically effective and cost-effective application of specific technologies and disease management strategies. In this context, the synthesis and critical appraisal of available evidence (assessment phase) is one input in a broader decision-making process that generates final recommendations (appraisal phase). NICE carries out its synthesis and critical appraisal functions in conjunction with other public entities.

The UK's Department of Health funds the National Institute for Health Research (NIHR), which undertakes three core functions that feed into NICE processes: (1) horizon scanning to inform the topic selection process; (2) evidence synthesis, including a systematic review of the evidence and decision analysis modeling or a critique of manufacturers' models, to inform the development of guidance on the use of specific technologies; and (3) starting in 2007, prospective real-world trials to address specific uncertainties identified during the guidance development process to inform future updates of the guidance "direct access."²¹

The Medical Research Council receives public funding to support research into methodologies for developing NICE guidance, including modeling tools for making conditional coverage decisions and ways for incorporating equity considerations into the decision-making algorithm. At the same time, NICE contracts with the Royal Colleges to synthesize and appraise the evidence.

The various sources of evidence reviewed during the assessment phase are synthesized and used to develop a decision analysis model. NICE has developed explicit guidance as to the type, format, and sources of evidence that decision-making committees consider during the appraisal process.²² The evidence includes meta-analyses and systematic reviews of RCTs, head-to-head RCT comparisons of the technologies under consideration, and different types of quasi-experimental studies, such as prospective cohort, registries, and epidemiological analyses.

Unpublished academic or proprietary commercial information is also eligible for consideration, although NICE encourages stakeholders to keep such submissions to a minimum, requires full justification of the confidential nature of such evidence, and expects the evidence to be made public following the licensing of a technology or publication of an academic analysis. Finally, NICE considers patient surveys and patient and professional expert opinion.

The appraisal phase—rather than the assessment phase—typically captures professional opinion while the NICE Citizens’ Council—a citizens’ jury consisting of 30 laypersons—elicits societal views on issues such as the importance of equity or age in making resource allocation decisions. The views of this citizens’ jury are captured and published in the NICE guideline on Social Value Judgments.²³

Often, CE entities are asked to make recommendations and issue guidance based on minimal publicly available information. In January 2007, the Citizens’ Council was asked to consider the role of decision makers such as NICE in encouraging the generation of CE evidence and reducing uncertainty around the effects of treatment. After three days of deliberations that involved researchers from UK universities and the Cochrane Collaboration, ethics committee members and ethicists, patients who had participated in clinical trials, researchers, industry representatives, and public payers, the council unanimously endorsed the idea of conditional coverage subject to further evidence. In its report, the council identified several criteria that should guide decision makers in making recommendations for additional research:

- The extent to which the research is likely to reduce current uncertainty;
- The value-for-money of the research;
- The budgetary implications of making a positive recommendation based on the evidence;
- The existence of an ongoing study or the feasibility of initiating a study within a realistic time frame; and
- Issues of patient access to the study across geographic areas.

In discussing the need for explicit criteria when making conditional coverage decisions for more CE and considering the difficulties of implementing needed

CE research, the council report concluded that “patients would be reassured to know that clinicians and the healthcare system in general could face up to uncertainty, and were confident enough to deal with it in a mature, scientific way, and avoid wasting money on unproven technologies.”

Drawing on the report, NICE asked its advisory bodies (Social Value Judgments)²⁴ to consider recommending the use of an emerging technology only in the context of a well-designed research project when limited evidence is available to support the technology’s broad adoption across the NHS.

NICE’s overall approach to evidence evaluation does not always follow conventional evidential hierarchies. Instead, it is driven largely by the type of policy and clinical practice questions that need to be answered and focuses on the quality (rather than the type) of the studies that address such questions. Further, given increasing pressures for making timely decisions, evidential synthesis rather than primary research often provides the impetus for the NICE approach, thereby allowing for:

- Consideration of several sources of evidence;
- Extrapolation beyond the usually short time horizons of most RCTs;
- Incorporation of epidemiological data specific to the UK population such as baseline risk or usual treatment patterns;
- Consideration of alternative technologies and their associated costs; and
- Quantification of uncertainty and the implications of making the wrong decision.

Increasingly, NICE is experimenting with decision options that link policy and practice recommendations to evidence

generation.²⁵ Risk-sharing schemes or conditional reimbursement decisions (similar to CMS’s Coverage with Evidence Development) are particularly relevant in circumstances of increased uncertainty as in the case of new drugs at the time of receiving marketing authorization or diagnostic tests and surgical procedures usually accompanied by limited evidence of impact on health outcomes.²⁶

Consideration of costs. *“Even if NHS funding is significantly increased that single truth will remain...resources do not stretch to satisfying the demands placed on them by everyone. No healthcare system in the world begins to meet, and match, the aspirations of all those who work in it or use it.”²⁷*

NICE considers costs from the NHS perspective (and, in exceptional circumstances, in terms of broader indirect societal costs) by requesting the calculation of an incremental cost-effectiveness ratio (ICER) that is measured as the cost per quality adjusted life year (QALY). Before reaching a decision, the NICE decision-making committees consider the ratio in addition to other non-quantifiable factors, such as equity implications or the availability of alternative treatments. The threshold range applied by NICE to assess the appropriateness of a technology’s ICER is between \$30,000 and \$45,000 per QALY, although advisory bodies may consider and have accepted ICERs out of range. In response to the negative reaction triggered by a provisional NICE decision to restrict access to four renal cancer drugs that extend life by an average few months at more than \$40,000 per patient per year, NICE has proposed to revise its threshold for end-of-life treatments. If approved, the proposal will make it easier for treatments targeting terminally ill patients to make it past NICE’s cost-effectiveness threshold.²⁸ The current threshold was initially based on anecdotal evidence from technology adoption decisions made by local purchasers. More recently, program budgeting and marginal analysis data released by the Department of Health made

it possible to estimate the substitution rate (or NHS productivity) at the local level. According to the analyses, the NICE threshold range is broadly compatible with NHS's return on investment in terms of reduction in mortality and (less easily measured) morbidity in the major disease areas such as cancer and cardiovascular disease.²⁹

NICE does not directly consider affordability but does provide budgetary impact analyses for all recommendations in order to support implementation. It also directs its committees to exercise greater caution and expect higher levels of certainty around the expected ICER when the budgetary implications of a recommendation are significant.

Germany

Topic selection. JFC, which is responsible for prioritizing topics referred to it by various professional associations, providers, and insurance funds, refers the vast majority of topics to IQWiG. About 10 percent of topics come directly from the German Ministry of Health and 2 to 3 percent from IQWiG. The pattern of referrals is consistent with IQWiG's broader model of operation, which serves as a "decision support tool" for JFC, with JFC stakeholders providing the large share of its financial support and representing the primary users of its advice. IQWiG is primarily responsible for selecting the topics for "information for patients" publications.

Assessment and appraisal. IQWiG is responsible for the assessment part of the decision-making process. As opposed to NICE, which hosts the decision-making committees, JFC and local providers make final listing and reimbursement decisions.

IQWiG operates with clear rules concerning the nature of evidence deemed appropriate to inform its advice to JFC. RCTs sit at the top of the evidential hierarchy. IQWiG specialists seek access to individual trial protocols and information on unpublished trials, especially for industry-sponsored

studies and, where available, patient-level data. They rarely adapt evidence syntheses undertaken in other countries to the German context and do not consider any confidential commercial information; a requirement mandates that all unpublished data considered by IQWiG must be made public. The limits on acceptable evidence often trigger complaints by industry and physician-proponents that the evidence base is inadequate.

Industry has also complained about its lack of involvement in evidence assessment. In response, IQWiG now permits industry comment at the various stages of the assessment, including the early priority-setting and preliminary report stages.

With regard to prospective trials, the law allows IQWiG to sponsor prospective trials to assess the comparative effectiveness of medical devices (rather than drugs), reflecting Germany's relatively more relaxed regulatory requirements for devices than drugs. In addition, Germany has developed a strong medical devices industry, although most of the device companies operating in Germany have limited financial means to sponsor RCTs. Similarly, IQWiG has not yet sponsored clinical trials of such devices largely because of the lack of dedicated funding.

JFC is the major customer for IQWiG's advice. Traditionally, JFC's decision-making process has been consensus-driven with relatively opaque decision rules that make it difficult to assess the impact of IQWiG's advice on the final decision. As mentioned, JFC is not obligated to accept IQWiG's advice but must discuss its rationale for not accepting such advice. The general lack of transparency at the appraisal and negotiation stage is now being addressed by opening JFC meetings to the public and encouraging the publication of IQWiG's evidence syntheses and the rationale for IQWiG's advice at the assessment stage.

Consideration of costs. As of 2007, IQWiG's scope of responsibility extends to costs. For many years, Germany used a national system for reference pricing for pharmaceuticals with generic equivalents. However, increasing pressures from expensive new drugs and the burgeoning costs of long-term care led to the introduction of a more central role for IQWiG in assessing cost-effectiveness. Starting in 2009, IQWiG will be responsible for advising insurance funds of a ceiling price for new and existing pharmaceuticals based on evidence of comparative cost-effectiveness. When the ceiling price is below the manufacturer's list price, the manufacturer will have to lower its price; if not, insurance funds will shift the extra cost to patients through co-payments.

IQWiG has adopted a unique approach to considering costs. As detailed in the law expanding IQWiG's responsibilities, IQWiG is not charged with considering the opportunity cost of coverage decisions across different technologies or disease areas but only within those. So, for example, it does not need to compare the effectiveness of cancer drugs to the effectiveness of drugs for hypertension but instead the cost effectiveness of an antihypertensive drug against other drugs to treat the same condition. Furthermore, only technologies with proven superiority to comparable technologies are subject to comparative evaluation of their costs. Individual insurance funds then use the ceiling price proposed by IQWiG to define the "appropriate" level of reimbursement, subject to JFC decisions.³⁰ In fact, the intent is to encourage individual insurance funds to negotiate prices down on a case-by-case basis, well below IQWiG's ceiling price, through direct negotiation with manufacturers in an environment of increased competition between funds.

Excluding consideration of opportunity costs and avoiding comparisons across diseases conveniently eliminates the need for an explicit threshold for quantifying benefits with application of a common utility measure (such as the QALY in the case of England and Australia).

Another concern is consideration of societal (indirect) costs such as productivity. Currently, IQWiG adopts a payer perspective (similar to the NICE model).

Perhaps the most innovative aspect of Germany's costing model is the adoption of an "efficiency frontier" approach for determining costs. This approach compares efficiency of individual treatments for a single indication to alternatives only after establishing their superior comparative clinical effectiveness. The upper price limit that the health care system should be prepared to pay for the treatments is then estimated in accordance with the unit cost set by the manufacturer and the efficiency (benefit/cost ratio) of existing treatments for the same indication. In a sense, the efficiency ratio of new (or existing) technologies is benchmarked against ratios for other technologies with the same indication rather than against a set threshold reflecting societal willingness to pay or affordability within a set budget. This approach avoids the controversial acknowledgment of finite resources and the recognition of opportunity costs. At the same time, it involves significant informational requirements... More important, it carries the risk of perpetuating inefficiencies for indications where current practice is wasteful. However, it may help contain costs by setting an upper price ceiling and may represent an important first step toward allowing decision makers to consider costs explicitly and consistently.

Some Unique Aspects of Methods to Inform Decisions in Other Countries

While most countries use some variation of topic selection, assessment and appraisal, and consideration of costs as implemented by NICE and IQWiG, some countries demonstrate unique approaches.

Ontario Health Technology Advisory Committee (OHTAC)

When conducting a first assessment of available evidence in the case of an insufficient basis for making an adoption (or rejection) decision, OHTAC

commissions its networks to undertake conditionally funded field evaluations (CoFFE).³¹ The networks include the Program for Assessment of Technology in Health (PATH) based at McMaster University and St. Joseph's Health Centre, the Toronto Health Economics and Technology Assessment (THETA) based at the University of Toronto, and the Institute for Clinical and Evaluative Sciences (ICES). With a budget of approximately \$8-10 million per annum (which in most cases does not include the cost of the new technologies³²), OHTAC has thus far (2008) completed more than 10 field evaluations with another 25 studies underway, including a combination of registries and RCTs for evaluating drug-eluting stents, a cardiac PET and an implantable cardiac defibrillator (ICD) registry, and two PET RCTs for staging early lung and breast cancer.³³

The OHTAC model is one of the few international examples of CE research in the form of prospective trials, as opposed to evidence synthesis and claims' data analysis, that is directly linked to decision making. OHTAC and the Medical Advisory Secretariat (MAS) call for research through the "field evaluation" model (similarly to CMS's coverage with evidence development or NICE's "only in research" scheme). What makes the Ontario model unique is that decision makers (not the National Institutes of Health or NIHR in the case of the United States or UK, respectively) control the research budget, thus making implementation of calls for research much more feasible.

Ontario's Institute for Clinical Evaluative Sciences (ICES)

ICES (www.ices.on.ca) was established in 1992 to "provide unique scientific insights to help policymakers, managers, planners, practitioners and other researchers shape the future direction of Ontario's health care system." Funded by Ontario's Ministry of Health (at approximately \$5 million per annum), ICES staff may receive grants from federal funding agencies, such as the

Canadian Institutes of Health Research. ICES's main focus is on undertaking population-level studies by using claims and administrative data through linked databases of individual patient information (up to 12 million Ontarians) instead of relying mostly on meta-analyses of RCTs (IQWiG) or decision analytic modeling (NICE). With a staff of over 100 scientists, ICES's objectives are to:

- Carry out population-based health services evaluation relevant to clinical practice and health policy development;
- Document province-wide patterns and trends in health care delivery;
- Develop and share evidence to inform decision making by policymakers, managers, clinicians, planners, and consumers;
- Promote linkages among health services researchers and decision makers; and
- Train researchers and promote a wider understanding of clinical epidemiology and health services evaluation.

In addition to conducting epidemiological studies to assess health trends and identify high-priority diseases in Ontario, ICES is commissioned by Ontario's Ministry of Health to evaluate policy initiatives, such as the impact of interventions to reduce waiting times for elective surgery for high-priority procedures (e.g., cataract or knee replacement surgery) on procedures not included on the priority list. ICES also provides information to OHTAC. At the request of OHTAC, ICES has set up an Ontario-wide Web-based registry to monitor the use and effectiveness of ICDs. ICD use data are collected in real time and may be linked to individual patient characteristics and outcomes at the patient level through ICES's large administrative databases. The results of analyses, published in peer-reviewed journals, inform province-wide coverage decisions.

Health Insurance Review Agency (HIRA)—Republic of Korea

In 2007, Korea abandoned reference pricing and switched to a positive drug listing system that calls for price negotiation based on evidence of comparative clinical and cost-effectiveness (value-based pricing). The shift highlighted the need to link decision making to evidence generation. In response, Korean policymakers introduced a system of conditional coverage for orphan drugs and inpatient cancer drugs based on what they termed weak evidence of CE. Furthermore, to support HIRA's decision-making role, the Korean parliament enacted a law in February 2008 to establish a national medical research institute (NECA) charged with undertaking research to address HIRA needs. Preliminary discussions indicate decision-maker interest in both head-to-head trials and claims analyses as well as in methods research, including social value elicitation from the Korean population and adaptation of QALY or another outcome measure to the Korean setting.

Pharmaceutical Benefits Advisory Committee (PBAC)—Australia

In Australia, drug companies nominate the price they are seeking. Based on comparative clinical and cost-effectiveness evidence, PBAC then decides whether the price is acceptable. In the absence of evidence of increased clinical benefit, price negotiation does not occur such that the analysis is a cost-minimization analysis with the “equieffective” dose as the primary issue. If the drug is cost-effective at the proposed price, then no price negotiation is needed; rather, negotiation focuses on risk-sharing schemes or price-volume agreements, especially in the case of considerable uncertainty around the clinical effect and any risk of “indication creep.”

A positive recommendation by PBAC usually means that a drug is listed, although listing always occurs at the prerogative of the Minister for Health and Ageing. However, the Minister may not reverse a negative PBAC recommendation. In other words, whereas a “yes” by PBAC is

a “maybe,” a “no” is always a “no.” In only one instance have political leaders ignored PBAC advice. In 2001, PBAC rejected Herceptin for breast cancer as not cost-effective. The Minister disagreed and, to bypass PBAC's recommendation and list Herceptin, he created a new program, the ongoing Herceptin program, which is outside PBAC's jurisdiction. This direct political intervention challenged PBAC's independence and led to a strong response.

PBAC's deliberations are informed by a detailed appraisal of the applicant's submission and the evidence contained within it, as prepared by four evaluation groups based at Australian universities, and by additional advice provided by specialist committees that take into account economic considerations and utilization patterns (Economics and Drug Utilization Subcommittees).³⁴ Occasionally, the Department prepares the needed analyses. The evaluation groups generate evidence consisting mainly of evidence synthesis, “model-busting” of manufacturers' models, and budgetary impact analyses. PBAC cannot recommend either conditional coverage of drugs or further research; it can, however, propose post-listing follow-up, especially when the industry submission is based on surrogate endpoints. It can also recommend, through the pricing authority, the establishment of a risk-sharing scheme, usually in the form of a price-volume agreement between government and the manufacturer.

Recent years have seen several developments in how the Australian system operates, including a commitment to increasing public involvement in decision making; an open call for submissions by interested parties for all drug evaluations to be considered by PBAC alongside the evaluation commentary; the introduction of cost-recovery arrangements to support PBAC processes by levying service fees on submitting companies; public disclosure of PBAC deliberations (in the form of Public Summary Documents) in response to the recent Free Trade Agreement between Australia and the United States; and an option for an independent review as an

alternative to resubmission when the PBAC declines to recommend the listing of a new drug (or new indication for an already listed drug). The independent review is not an appeal mechanism but rather, considers issues in dispute and cannot repeal a PBAC recommendation. Moreover, the outcome of a review does not mandate PBAC's reconsideration of an application.

High Health Authority—France

HAS plays a central role in pricing new medical technologies. Its specialized committees review the absolute and comparative clinical effectiveness of new technologies and then rank them on a 5-point scale, from no improvement to significant improvement. Then, a separate entity determines the price premium justified by the clinical improvement.

Since 2008, HAS is required by law to consider costs when issuing its advice. Economic considerations are included in those cases where HAS considers classes of drugs or other technologies rather than individual products; accordingly, the introduction of economic evaluation is not expected to affect pricing decisions.

Pharmaceutical Benefits Board (LFN)—Sweden

Sweden's Pharmaceutical Benefits Board (LFN) was established in 2002 and is responsible for developing and maintaining a national positive list for drugs whose costs are reimbursed by the health care system. In making its listing decisions, LFN considers comparative clinical and cost-effectiveness, clinical need (severity of disease), and equality. In its economic evaluation, LFN adopts a societal (rather than payer) perspective, including, where appropriate, factors such as productivity costs.

LFN decisions, in conjunction with mandatory generic substitutions for branded drugs (otherwise the patient pays the difference out of pocket), have prompted many drug companies to lower their prices on branded drugs. Manufacturers want to maintain their drugs

in the positive list. Generic substitution and LFN CE reviews of classes of drugs, such as proton pump inhibitors, migraine drugs, and antihypertensives, have resulted in approximately \$700 billion in savings in recent years.

Dissemination and Impact

CE entities need to be able to communicate their findings to a wide audience. Countries rely on a variety of mechanisms to disseminate their finding and evaluate the impact of their recommendations.

England (UK)

The initial establishment of NICE explicitly excluded implementation from the institute's responsibilities. However, unwarranted geographic variation persisted despite national guidance. As a result, the implementation of NICE recommendations became a government priority. Among the measures introduced to reduce regional variation was the requirement that primary care trusts make funds available to implement NICE decisions. The recent announcement that NHS patients will (subject to enactment of the necessary law) be entitled to access new technologies approved by NICE in the context of the NHS Constitution³⁵ is expected to increase uptake of NICE recommendations and reduce geographic variation.

NICE's implementation directorate has developed several tools and interventions for supporting the adoption of NICE guidance at the local level, including audit criteria, educational tools, a network of "implementation consultants" operating at the local level, guides to changing provider behavior, budget impact tools adaptable to the local setting, and a "forward planner" to help commissioners plan for NICE guidance in the pipeline. The implementation directorate is also responsible for collating and sharing case studies of best practice in implementing NICE guidance and for developing and

maintaining a database of uptake studies from across the UK (www.nice.org.uk/ernie).

The evaluation of the impact of NICE is partially based on ideology and partially based on indirect measures of impact. For example, there is concern that NICE slows the diffusion of new technologies and recognition of the occasional difficulty in linking clinical and economic evaluations given different evaluation methods. Notwithstanding these concerns, most evaluations are positive.

According to the latest assessment of NIHR's health technology assessment program, 96 percent of academic groups providing evidence for consideration by NICE committees reported an impact on policy versus 60 percent of those undertaking other NIHR-funded research.³⁶ Researchers who want their research to have an impact are apparently satisfied.

It is methodologically challenging to assess the impact of NICE guidance on practice patterns and even more so on health outcomes. The lack of a control group and the wide range of government policies and non-health-related activities make it impossible to attribute causality to NICE. However, numerous case studies show the impact of the use of CE to inform coverage decisions on unwarranted variation in practice and spur the diffusion of new treatment modalities across the NHS. According to a report by the National Director for Cancer, adoption of cancer drugs appraised by NICE increased by almost 50 percent across the country between 2003 and 2005, with geographic variation declining from three- to eight-fold to two- to three-fold over the same period.³⁷

NICE was not intended to generate cost savings when it was established. Adoption decisions have been estimated to cost \$1.6 to \$3.2 billion per year. In 2006–2007, NICE guidance absorbed more than a tenth of the growth in health care spending across the NHS.

Perhaps NICE's greatest contribution has been an increased awareness of the importance of evidence-informed health care resource allocation decisions in a transparent, inclusive, and methodologically robust way. In its 2008 parliamentary enquiry into NICE, the bipartisan Health Select Committee concluded, "NICE does a vital job in difficult circumstances. The development of more and more health technologies and procedures, alongside rising patient expectations and the ageing population, is going to make it even more difficult in the future. Healthcare budgets in England, as in other countries, are limited. Patients cannot expect to receive every possible treatment. Demand outstrips resources and priorities have to be determined. In other words rationing is essential, and NICE has a key role to play. Given the difficult environment, NICE requires the backing of the Government. NICE must not be left to fight a lone battle to support cost- and clinical effectiveness in the NHS."³⁸

Germany

Implementation of IQWiG's recommendations has been difficult to evaluate given (1) the strictly advisory nature of IQWiG's guidance as defined by law and (2) the lack of explicitness in justifying JFC coverage decisions in light of IQWiG's advice. In the absence of a mandate for JFC to adopt IQWiG's recommendations, the requirement for JFC to explain its grounds for rejection, coupled with the recently announced opening of its decision-making meetings to the public, will increase the pressure on JFC to rely on IQWiG's evidence or produce convincing reasons for reaching an alternative conclusion.

The introduction of an IQWiG-set price ceiling for medical technologies is also expected to increase IQWiG's impact on industry pricing and physician prescribing practices. However, the German model of delegating decision-making power to insurance funds means that IQWiG has an indirect and difficult-to-assess impact

on decisions at the national level. When JFC accepts IQWiG's advice, the advice tends to be binding for insurance funds. However, the link between IQWiG's advice and JFC decisions remains unclear.

Australia and Canada

PBS in Australia and CDR in Canada make listing decisions for drugs for the whole of their respective populations (in the case of CDR, for public plans only). Consideration and approval by PBAC is a prerequisite for a drug's inclusion in the Australian formulary while, even though state-run public plans in Canada operate with a degree of autonomy, the vast majority of plans follow CDR's recommendations. In this sense, both entities have considerable impact. However, their scope is limited to pharmaceuticals and, in the case of CDR,

to public payers only (employer-sponsored plans make their own listing decisions in Canada).

Acknowledgments

The authors are grateful to the following experts who consulted on the role and functions of:

- OHTAC and CDR: Irfan Dhalla, Les Levine, Danielle Whicher
- PBS: Ruth Lopert
- HAS: Lise Rochaix
- IQWiG: Charalbos-Markos Dintsios, Jonas Schreyoegg

- NICE: Sir Michael Rawlins, Andrew Dillon

As well as to the many individuals who provided insight and feedback on the manuscript, including:

- Robert Crane, Kaiser Permanente
- Amanda Folsom, AcademyHealth
- Patricia Pittman, AcademyHealth

The views expressed herein are those of the authors not of their employers, the experts consulted or of the sponsors of this analysis.

This report is part of a three year initiative to expand senior health policy leaders' knowledge of health systems and delivery innovations in other industrialized countries, and to promote the use of this knowledge to advance system improvements in the United States. AcademyHealth gratefully acknowledges the Kaiser Permanente Institute for Health Policy and the National Institute for Health Care Management Foundation for their generous support of this work. See www.academyhealth.org for more information.



Appendix 1: Defining Comparative Effectiveness

Table 1 depicts some alternative definitions of CE as proposed by various organizations in the United States prior to the release of the Federal Coordinating Council for Comparative Effectiveness Research. Despite considerable overlap in the definitions, the table indicates significant differences, which were indicative of some of the debate over the role of CE.

Table 1: Definitions of Comparative Effectiveness by Various U.S. Organizations

Organization	Definition
American College of Physicians	“Comparative effectiveness is the evaluation of the relative (clinical) effectiveness, safety, and cost of 2 or more medical services, drugs, devices, therapies, or procedures used to treat the same condition.” ³⁹
Institute of Medicine	“. . .comparative effectiveness [is] the comparison of one diagnostic or treatment option to one or more others. In this respect, primary comparative effectiveness research involves the direct generation of clinical information on the relative merits or outcomes of one intervention in comparison to one or more others, and secondary comparative effectiveness research involves the synthesis of primary studies to allow conclusions to be drawn.” ⁴⁰
Medicare Payment Advisory Committee	“Comparative-effectiveness analysis compares the relative value of drugs, devices, diagnostic and surgical procedures, diagnostic tests, and medical services. By value, we mean the clinical effectiveness of a service compared with its alternatives. Comparative-effectiveness information has the potential to promote care of higher value and quality in the public and private sectors.” ⁴¹
Congressional Budget Office	“As applied in the health care sector, an analysis of comparative effectiveness is simply a rigorous evaluation of the impact of different options that are available for treating a given medical condition for a particular set of patients.” ⁴²
Agency for Healthcare Research and Quality	“Comparative effectiveness reviews expand the scope of a typical systematic review, which focuses on the effectiveness of a single intervention, by comparing the relative benefits and harms among a range of available treatments or interventions for a given condition. In doing so, [they] more closely parallel the decisions facing clinicians, patients and policymakers, who must choose among a variety of alternatives in making diagnostic, treatment, and health care delivery decisions.” ⁴³
Pharmaceutical Research and Manufacturers of America	“. . .the development of high quality [sic] information about patients’ medical treatment options.” ⁴⁴

Appendix 2: Comparative Effectiveness Research Infrastructure

Even with some agreement regarding the definition of comparative effectiveness, differences in the preferred organization and domains of the CE entity were prevalent prior to the passage of the ARRA legislation. For example, various organizations in the United States had proposed a wide range of alternative organizational forms with a variety of public and private roles (Table 2).

Table 2: Organizational Forms for Comparative Effectiveness

Organization	Proposal
America's Health Insurance Plans	Establish a national (public/private) entity to evaluate new and existing health care services and technologies. The new entity (which could be known as a Comparative Effectiveness Board) would be responsible for (1) comparing the clinical and cost-effectiveness of new and existing drugs, devices, procedures, therapies, and other health care services; (2) assessing alternative uses of treatments currently in practice; and (3) disseminating the resultant information in a useful format, allowing patients and clinicians to make informed health care decisions. ⁴⁵
The Commonwealth Fund	Establish a Center for Medical Effectiveness and Health Care Decision-Making to improve decision making and incorporate information about the relative clinical benefits and cost-effectiveness of alternative treatment options into insurance benefit design. By generating the information and creating payment and cost-sharing incentives for providers and consumers, this policy option could result in estimated health system savings of \$368 billion shared over 10 years by all payers. ⁴⁶
Institute of Medicine	Congress should direct the U.S. Department of Health and Human Services to establish a program with the necessary authority, expertise, and resources to set priorities for evaluating clinical services and to conduct systematic reviews of the evidence. The program would develop and promote rigorous standards for creating clinical practice guidelines, which could help minimize use of questionable services and target services to patients most likely to benefit. ⁴⁷
Medicare Payment Advisory Commission	Congress should charge an independent entity to sponsor credible research on comparative effectiveness of health care services and disseminate this information to patients, providers, and public and private payers. Such an entity would be independent, transparent, and inclusive, with secure funding, and would be responsible for regular review, dissemination of its findings to relevant stakeholders and would be overseen by an independent Board. ⁴⁸
Medicare Coverage Advisory Committee	The Centers for Medicare and Medicaid Services launched the Evidentiary Priorities Initiative in 2007 by convening a panel of the Medicare Evidence Development & Coverage Advisory Committee (MedCAC) to identify high-priority topics for CE. ⁴⁹
Congressional Budget Office	Given that any private sector entity (such as a health plan) has only a limited incentive to produce or pay for information that could benefit many entities—including its competitors—it is reasonable to argue for a larger federal role in coordinating and funding research on comparative effectiveness. In addition, with federal health insurance programs playing a major role in financing medical care and accounting for a large share of the federal budget, the federal government has an interest in generating evaluations of the effectiveness of different approaches to health care. ⁵⁰
Congressional Bills	As just one example, the Conrad/Baucus Bill (August 2008), which is entitled the Comparative Effectiveness Research Act of 2008, recommends the establishment of a comparative effectiveness research fund to prioritize and undertake CE research.
American College of Physicians	A 2008 position statement ⁵¹ called for an “adequately funded, trusted national entity to prioritize, sponsor, and produce comparative-effectiveness and cost-effectiveness information.”

In addition, CE has been the topic of a series of academic publications⁵² and policy reports⁵³ developed by U.S. health research and policy organizations, including AcademyHealth.⁵⁴

Endnotes

- Office of Technology Assessment, Congress of the United States, "Identifying Health Technologies That Work: Searching for the Evidence," September 1994.
- For example, the decision of the Australian Minister to ignore the negative recommendation by PBAC on Herceptin resulted in a series of resignations of committee members while a similar decision by the UK government to ignore NICE's recommendation on drugs for multiple sclerosis has been one of the most controversial decisions in the organization's history.
- Department of Health, "The New NHS: Modern, Dependable," December 1997, <http://tinyurl.com/4asgpy>.
- For a detailed description of NICE and its functions, see Sorenson, C. et al., "NICE: How Does It Work and What Are the Implications for the US?" Funded by the National Pharmaceutical Council, 2008.
- Ibid.*, DH, The New NHS, paragraph 7.6, 1997.
- Cheng, M. and U. Reinhardt, "Shepherding Major Health System Reforms: A Conversation with German Health Minister Ulla Schmidt," *Health Affairs*, Vol. 27, No. 3 (2008):w204-w213. There are no pay-for-performance schemes in Germany. Sickness funds are paid according to a risk-adjusted per capita formula, and sickness funds that produce better outcomes are not rewarded.
- Sloan, C. "A History of the Pharmaceutical Benefits Scheme 1947-1992," Commonwealth Department of Human Services and Health. Canberra, Australia, 1995.
- Speech by the Minister of Labour and National Service of Australia, 1944.
- Lopert, R. "Pharmacoeconomics and Drug Subsidy in Australia—An Account of the Australian Approach. In Patients Physicians and Pharma Workshop Report, Israel National Institute for Health Policy Research, Haifa, April 2006, <http://www.israelhpr.org.il/FileServer/3434b707e1c2e288bbd73a2f1bf941f.pdf>; also see National Health Act 1953, S101.
- Rochaix, L. and D. Wilsford. "State Autonomy, Policy Paralysis: Paradoxes of Institutions and Culture in the French Health Care System." *Journal of Health Politics, Policy and Law*, Vol. 30, No. 1-2 (Feb.-April 2005).
- Conversations at the Ministry of Health Brasilia, Brazil, with Dr. Anderson.
- Department of Health, "Directions to Primary Care Trusts and NHS Trusts in England Concerning Arrangements for the Funding of Technology Appraisal Guidance from NICE," July 1, 2003, <http://tinyurl.com/3guja6>.
- Department of Health, "The National Health Service Constitution," June 30, 2008, http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_085814.
- Department of Health, "Research for Health: A Research and Development Strategy for the NHS," London, 1991.
- See <http://www.mrc.ac.uk/NewsViewsAndEvents/News/MRC004215>.
- Gemeinsamer Bundesausschuss, "The German Health Care System and the Federal Joint Committee," <http://www.g-ba.de/institution/sys/english/>.
- Healthcare Modernization Act of 2004, paragraph 139a.
- In Canada, general hospitals are nonprofit entities operating on budgets negotiated with provincial Ministries of Health with limited capacity (or incentive) to invest in new technologies due to financial pressures. These attributes of the demand side may explain the establishment and buy-in to the OHTAC model in Ontario.
- Topic selection criteria include broad policy priorities, potential budgetary impact, potential to improve health outcomes, current variation in practice, availability of relevant evidence, and potential of NICE guidance to add value.
- National Institute for Health and Clinical Excellence, "NICE Topic Selection Process," http://www.nice.org.uk/getinvolved/suggestatopic/topicselectionconsiderationpanels/topic_selection_consideration_panels.jsp, accessed October 2008.
- Chalkidou, K. et al., "Evidence-Informed Evidence-Making," *Journal of Health Services, Research & Policy*, Vol. 13, No. 3 (2008):167-73.
- Guide to the Methods of NICE Technology Appraisals, NICE, 2008; also see Claxton, K. et al., "Probabilistic Sensitivity Analysis for NICE Technology Assessment: Not an Optional Extra," *Health Economics*, Vol. 14, No. 4 (2005):339-47.
- National Institute for Health and Clinical Excellence, "Social Value Judgments: Principles for the Development of NICE Guidance," 2005-2008, http://www.nice.org.uk/getinvolved/patientandpublicinvolvement/opportunitiestogetinvolved/citizenscouncil/socialvaluejudgementsguidelines/social_value_judgements_guidelines.jsp.
- See above
- Chalkidou, K., et al., "Making a Decision to Wait for More Evidence: When NICE Recommends a Technology Only in the Context of Research," *Journal of the Royal Society of Medicine* 100 (2007):453-60; K. Chalkidou et al., "Evidence-Based Decision Making: When Should We Wait for More Information?" *Health Affairs*, Vol. 27, No. 6, in press.
- Tunis, S. and S. Pearson, "Coverage Options For Promising Technologies: Medicare's 'Coverage With Evidence Development,'" *Health Affairs*, Vol. 25, No. 5 (2006): 1218-1230; Chalkidou, K. and J. Lord, et al., "Evidence based decision making in healthcare: when should we wait for more information?" *Health Affairs*, Vol. 27, No. 6 (2008): 1642-1653
- National Institute for Health and Clinical Excellence, "Speech to St. Paul Healthcare by Professor Sir Michael Rawlins," September 7, 1999, http://www.nice.org.uk/aboutnice/whatwedoniceandthenhs/speech_to_st_paul_healthcare_by_professor_sir_michael_rawlins.jsp.
- National Institute for Health and Clinical Excellence, "Appraising End-of-Life Medicines: A NICE Consultation," November 2008, <http://www.nice.org.uk/newsevents/infocus/EndOfLifeMedicinesConsultation.jsp>.
- Martin, S., et al., "The Link between Healthcare Spending and Health Outcomes: Evidence from English Programme Budgeting Data," June 2007, http://www.health.org.uk/publications/research_reports/the_link_between.html and follow-on May 2008, http://www.health.org.uk/publications/research_reports/spending_outcomes.html.
- However, in reality, individual sickness funds have little power to amend JFC decisions, which tend to be binding.
- Levin, L., et al., "Establishing a Comprehensive Continuum from an Evidentiary Base to Policy Development for Health Technologies: The Ontario Experience," *International Journal of Technology Assessment in Health Care*, Vol. 23, No. 3 (2007):299-309.
- The cost of the new technologies is also usually covered by the Ministry through a separate fund, thus making implementation of field evaluations a relatively appealing option for manufacturers, especially if the alternative is no adoption by local providers.
- Whicher, D., et al. Comparative Effectiveness Research in Ontario, Canada: Producing Relevant and Timely Information for Health Care Decision Makers Milbank Quarterly, fall 2009 issue (*in press*)
- Affordability is a key consideration for PBAC, in addition to comparative cost-effectiveness; for approved technologies with an estimated budget impact of more than AUS\$10 million, the Cabinet (rather than the Minister) grants final approval.
- Ibid.*, NHS Constitution.
- Hanney S., et al. "An Assessment of the Impact of the NHS Health Technology Assessment Programme," *Health Technology Assessment*, Vol. 11, No. 53 (2007).
- Department of Health, "Usage of Cancer Drugs Approved by NICE," September 20, 2006, http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_4139061.
- The House of Commons Health Select Committee, "First Report, Session 2007-08: National Institute for Health and Clinical Excellence," January 10, 2008, <http://www.publications.parliament.uk/pa/cm/cmhealth.htm>.
- American College of Physicians, "Information on Cost-Effectiveness: An Essential Product of a National Comparative Effectiveness Program," *Annals of Internal Medicine* 148 (2008).
- Institute of Medicine Roundtable on Evidence Based Medicine, "Learning What Works Best: The Nation's Need for Evidence of Comparative Effectiveness in Healthcare," September 2007.
- Medicare Payment Advisory Commission, Report to Congress, "Promoting Greater Efficiency in Medicare," June 2007.
- Ibid.*, CBO, 2007.
- Agency for Healthcare Research and Quality, "Methods Reference Guide for Effectiveness and Comparative Effectiveness Reviews," October 2007.
- PhRMA, "Principles for Government-Supported Health Outcomes Research on Medical Technologies and Services," http://www.phrma.org/health_outcomes_research/.
- Ibid.*, AHIP, 2007; also see America's Health Insurance Plans, "A Shared Responsibility: Advancing Toward a More Accessible, Safe, and Affordable Health Care System for America," May 2008.
- Commonwealth Fund, "Bending the Curve: Options for Achieving Savings and Improving Value in the U.S. Health Spending," December, 2007.
- Institute of Medicine, "Knowing What Works in Healthcare: A Roadmap for the Nation," Washington, DC: National Academies Press, 2008, www.iom.edu/CMS/3809/34261/50718.aspx.

- 48 Medicare Payment Advisory Commission, "Producing Comparative-Effectiveness Information." In Report to the Congress, *Promoting Greater Efficiency in Medicare*, Washington, DC.; Medicare Payment Advisory Commission, 2007, www.medpac.gov/chapters/Jun07_Ch02.pdf.
- 49 Centers for Medicare and Medicaid Services, "Medicare Evidentiary Priorities," October 2007, http://www.cms.hhs.gov/CoverageGenInfo/07_EvidentiaryPriorities.asp, accessed April 2008.
- 50 Congress of the United States, Congressional Budget Office, "Research in the Comparative Effectiveness of Medical Treatments," December 2007.
- 51 Kirschner N. "Improving Availability of Comparative Effectiveness and Cost-Effectiveness Information: An Essential Feature for a High-Quality and Efficient United States Health Care System," A Position Paper of the American College of Physicians, Philadelphia: American College of Physicians, 2008.
- 52 Wilensky G.R. "Developing a Center for Comparative Effectiveness Information." *Health Affairs* 25 (2006) w572-85; E.J. Emanuel et al., "Essential Elements of a Technology and Outcomes Assessment Initiative," *Journal of the American Medical Association* 298 (2007):1323-5.
- 53 Congressional Research Service, Report to the U.S. Congress, "Comparative Clinical Effectiveness and Cost-Effectiveness Research: Background, History, and Overview," October 2007.
- 54 Holve E. and Pittman, P. "A First Look at Comparative Effectiveness Research in the United States," 2009, AcademyHealth.