

Health Policy Review



Facts, Fallacies, and Politics of Comparative Effectiveness Research: Part I. Basic Considerations

Laxmaiah Manchikanti, MD¹, Frank J.E. Falco, MD², Mark V. Boswell, MD, PhD³,
and Joshua A. Hirsch, MD⁴

From: ¹Pain Management Center of Paducah, Paducah, KY; ²Mid Atlantic Spine & Pain Specialists of Newark, Newark, DE; ³Texas Tech University Health Science Center, Lubbock, TX; and; ⁴Massachusetts General Hospital and Harvard Medical School, Boston, MA.

Dr. Manchikanti is Medical Director of the Pain Management Center of Paducah, Paducah, KY, and Associate Clinical Professor of Anesthesiology and Perioperative Medicine, University of Louisville, Louisville, KY.

Dr. Falco is Medical Director of the Mid Atlantic Spine & Pain Specialists of Newark, DE and Clinical Assistant Professor, Temple University Medical School, Philadelphia, PA.

Dr. Boswell is Professor of Anesthesiology and Director of the International Pain Center, Texas Tech University Health Sciences Center, Lubbock, TX.

Dr. Hirsch is Chief of Minimally Invasive Spine Surgery, Depts. of Radiology and Neurosurgery, Massachusetts General Hospital and Associate Professor of Radiology, Harvard Medical School, Boston, MA.

Address correspondence:
Laxmaiah Manchikanti, M.D.
2831 Lone Oak Road
Paducah, Kentucky 42003
E-mail: drlm@thepainmd.com

Disclaimer: There was no external funding in preparation of this manuscript. Conflict of Interest: Dr. Hirsch is a consultant for Cardinal Healthcare. He is a minor shareholder in Medtronic and Cardinal Healthcare. He serves on the Steering Committee for KAVIAR trial (volunteer position), and on the Data and Safety Monitoring Board (DSMB): CEEP trial (volunteer position).

Manuscript received: 01/04/2010
Accepted for publication: 01/12/2010

Free full manuscript:
www.painphysicianjournal.com

While the United States leads the world in many measures of health care innovation, it has been suggested that it lags behind many developed nations in a variety of health outcomes. It has also been stated that the United States continues to outspend all other Organisation for Economic Co-operation and Development (OECD) countries by a wide margin. Spending on health goods and services per person in the United States, in 2007, increased to \$7,290 – almost 2½ times the average of all OECD countries. Rising health care costs in the United States have been estimated to increase to 19.1% of gross domestic product (GDP) or \$4.4 trillion by 2018. The increases are illustrated in both public and private sectors.

Higher health care costs in the United States are implied from the variations in the medical care from area to area around the country, with almost 50% of medical care being not evidence-based, and finally as much as 30% of spending reflecting medical care of uncertain or questionable value. Thus, comparative effectiveness research (CER) has been touted by supporters with high expectations to resolve most ill effects of health care in the United States and provide high quality, less expensive, universal health care.

CER is defined as the generation and synthesis of evidence that compares the benefits and harms of alternate methods to prevent, diagnose, treat, and monitor a clinical condition or to improve the delivery of care. The efforts of CER in the United States date back to the late 1970's even though it was officially born with the Medicare Modernization Act (MMA) and has been rejuvenated with the American Recovery and Reinvestment Act (ARRA) of 2009 with an allocation of \$1.1 billion.

CER has been the basis for health care decision-making in many other countries. According to the International Network of Agencies for Health Technology Assessments (INAHTA), many industrialized countries have bodies that are charged with health technology assessments (HTAs) or comparative effectiveness studies. Of all the available agencies, the National Institute for Health and Clinical Excellence (NICE) of the United Kingdom is the most advanced, stable, and has provided significant evidence, though based on rigid and proscriptive economic and clinical formulas.

While CER is making a rapid surge in the United States, supporters and opponents are expressing their views. Part I of this comprehensive review will describe facts, fallacies, and politics of CER with discussions to understand basic concepts of CER.

Key words: Comparative effectiveness research, evidence-based medicine, Institute of Medicine, National Institute for Health and Clinical Excellence, interventional pain management, interventional techniques, geographic variations, inappropriate care.

Pain Physician 2010; 13:E23-E54

Despite a plethora of publications of comparative effectiveness in favor and against, over the past year, comparative effectiveness research (CER) has become the most celebrated research initiative in medicine in the United States. However, CER has been practiced too long to be considered new. Even then, it is too recent a concept to be considered standard practice. The intellectual roots of effectiveness research can be traced back to mid 18th century Scotland and the "arithmetical medicine" practiced by the graduates of Edinburgh Medical School (1). The first comparative effectiveness study was initiated by James Lind who undertook a controlled trial of 6 separate treatments for scurvy (2). In the United States, Ernest Codman, at the beginning of the 20th century, founded "outcomes management" in patient care (3). Even though its reviews of "comparative effectiveness" or CER are considered systematic, it builds on skepticism, the investigation of variations, randomized controlled trials (RCTs), and cost-benefit analysis.

CER is defined by the Institute of Medicine (IOM) (4) as, "The generation and synthesis of evidence that compares the benefits and harms of alternative methods to prevent, diagnose, treat, and monitor a clinical condition or to improve the delivery of care. In contrast, evidence-based medicine (EBM) is defined (5) as, "The conscientious, explicit, and judicious use of current best evidence in making decisions about the care of individual patients."

EBM is essentially focused upon the use of the right (types and extent of) knowledge to guide the right and good intentions and actions of medical practice, which is fundamental to prudential clinic decision-making (6,7). In contrast, CER is to assist consumers, clinicians, purchasers, and policy-makers to make informed decisions that will improve health care at both the individual and population levels. Thus, EBM and CER share many similarities and goals. They are analogous to religion and politics – meaning different things to different people (8-18).

CER, once only the scientific interest of clinical and health services researchers who compared medical treatments, now has become one of the hottest political health care issues. Facing the need for drastic improvement in our nation's health care delivery, Congress and the Obama Administration are looking to CER to improve and broaden the use of treatments in a cost-effective manner (4,8,9). It is passionately debated by proponents and opponents with equal comparative effectiveness evidence, or lack thereof (3,8). The

Obama Administration and Congress are enthusiastically moving forward with CER to improve and broaden the use of treatments in a cost-effective manner (19). Consequently, some researchers, clinicians, professional societies, and policy experts have welcomed this, as they see CER as a scientifically rigorous way to select the most effective treatments for the benefit of patients and the public (8). However, upon a closer look into various aspects of CER and its utilization in other countries, including National Institute for Health and Clinical Excellence (NICE) as part of the National Institute of Health (NIH) in England, apprehension has surfaced among the physician community and the public in general, along with pharmaceutical, biotech, and medical device companies (20-22). Those who pay for treatments, including health plans, insurers, and large employers foresee major changes in the way we practice medicine in the United States. According to opponents, this essentially translates to rationing of health care and removal of some or many treatments which we utilize. The proponents enumerate the potential benefits of cost effectiveness research, including scientific knowledge, improved health, and financial impact (16-18). Under the CER, while all medicine is impacted, the impact is highly variable based on specialties.

Interventional pain management is an evolving specialty. Interventional pain management encompasses the discipline of medicine devoted to the diagnosis and treatment of pain related disorders principally with the application of interventional techniques in managing sub acute, chronic, persistent, and intractable pain, independently or in conjunction with other modalities of treatment as a specialty designated as -09 in 2002 (23). The mainstay of interventional pain management is interventional techniques. They are minimally invasive procedures including percutaneous precision needle placement, with injection of drugs in targeted areas or ablation of targeted nerves; and some surgical techniques such as laser or endoscopic discectomy, intrathecal infusion pumps and spinal cord stimulators, for the diagnosis and management of chronic, persistent or intractable pain (24). Interventional pain physicians – rightfully so – are apprehensive about the impact of CER. This may be related to a new specialty or it may be related to involvement of many organizations and specialties with claims of ownership to the specialty. Thus, we will explore the role of CER in general and its impact on interventional pain management.

Part 1 of this review will discuss the basic aspects of CERs.

1.0 WHY CER?

The pace of innovation in health care has never been greater, and this innovation is constantly adding to broad and complex areas of health care interventions and systems (6,7,9-15). Thus, the need for careful scientific evaluation of clinical practice became a prominent focus during the second half of the 20th century (25). The demonstration of pervasive and persistent unexplained variability in clinical practice, and high rates of perceived inappropriate care, combined with increased expenditures, have fueled a steadily increasing demand for evidence of clinical effectiveness (6,7,9,26-83). Consequently, a body of evidence regarding safety, effectiveness, appropriate indications, cost-effectiveness, and other attributes of medical care are demanded. Failure to understand which services work best under what circumstances, and for which types of patients, contributes to the increasing cost of care, threats to patient safety, and avoidable loss of life (84).

The United States has the most expensive health care in the world by a large margin. However, it has been claimed that by many measures of public health, the United States ranks well down the list of nations (9). Thus, in spite of unprecedented advances in biomedical knowledge and the highest per capita health care expenditures in the world, the quality and outcomes of health care vary dramatically across the United States.

1.1 Rising Health Care Costs

Health at a Glance 2009, Organisation for Economic Co-operation and Development (OECD) indicators (85), estimates total expenditures on health measures considered as the final consumption of health care goods and services plus capital investment in health care infrastructure, which includes spending by both public and private sources (including households) on medical services and goods, public health and prevention programs and administration. Based on this report, in 2007, the United States continued to outspend all other OECD countries by a wide margin. Spending on health goods and services per person in the United States, in 2007, increased to \$7,290 – almost 2.5 times the average of all OECD countries. Most of the northern and western European countries, together with Canada and Australia, spent between \$3,000 and \$4,000, between 100% and 130% of the OECD average. However, Japan spent less on health than the average OECD countries, despite its above-average per capita income. Norway and Switzerland spent about two-thirds of the per capita level of the United States, but are still around 50% above the OECD average.

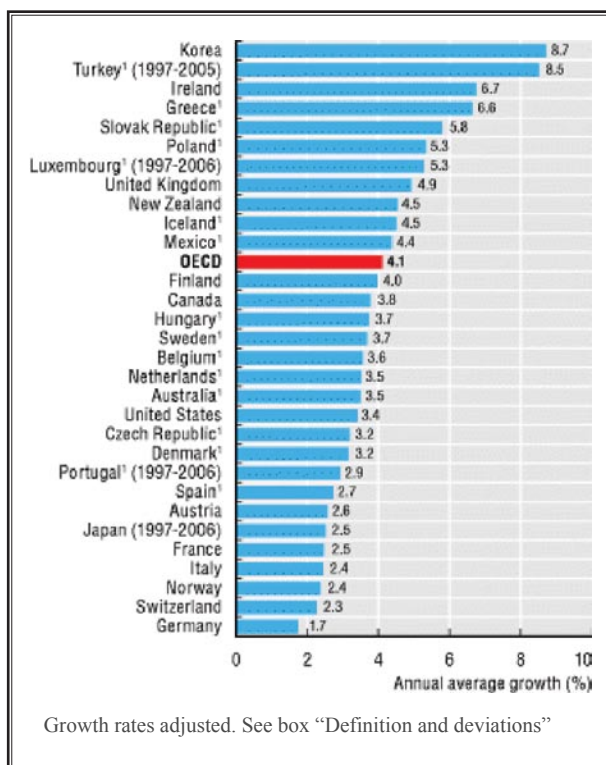


Fig. 1. Annual average real growth in per capita health expenditure, 1997-2007.

Source: OECD (2009), *Health at a Glance 2009: OECD Indicators*, OECD Publishing (85).

Figure 1 illustrates per capita spending on health with separation of public and private components. Overall, the variation in the levels of public spending on health is similar to that observed for total spending on health. Thus, it is estimated that even if the private sector in the United States continues to play the dominant role in financing, public spending on health per capita is still greater than that in most other OECD countries, because overall spending on health is much higher than in other countries. A large proportion of health care financing comes from private sources in Switzerland, whereas in Denmark, most health care is publicly financed as in many other countries.

Per capita health spending over 1997 to 2007 also is estimated to have grown, in real terms, by 4.1% annually on average across OECD (Fig. 1). In Germany, health spending per capita increased, in real terms, by 1.7% per year on average, the lowest of all the OECD nations, reflecting the effect of cost-containment policies designed to achieve stable contribution rates by

employers and employees. These measures have included budget or spending caps for sectors or individual providers, introducing reference prices for pharmaceuticals; educational approaches to enhance generic and rational prescribing; reducing the number of hospital beds; restricting the number of high cost medical equipment; and introducing or increasing co-payments for certain services. Other countries, such as Ireland and the United Kingdom, pursued specific policy objectives to increase public spending on health, resulting in overall health spending that outpaced economic growth. However, the real growth rate has been 3.4% in the United States, much below many other countries, but higher than a few countries, including Germany, Switzerland, and Japan.

In 2007, OECD countries devoted 8.9% of their gross domestic product (GDP) to health care spending (Fig. 2) (85). Trends in health spending to GDP ratio are

the result of the combined effect of trends in both GDP and health expenditures. In almost all countries, health spending grew more quickly than GDP over the last 10 years. It is expected that the share of health expenditure to GDP is likely to increase further, due to the recession that started in many countries in 2008. The share of health spending to GDP ranged from 11% in France, 10.4% in Germany, and 10.1% in Canada.

Current health expenditure in the U.S., as a share of household consumption in 2007, which is almost 13%, with the vast majority of OECD countries devoting more than 10% of their consumption to health with 5 countries, including the United States, Switzerland, Luxemburg, Norway, and Austria spending more than 15% on health in 2007 (85).

Health care costs in the United States have been estimated to increase to 19.1% of GDP or \$4,359 trillion by 2018 (Fig. 3) (86). The increases were forecast

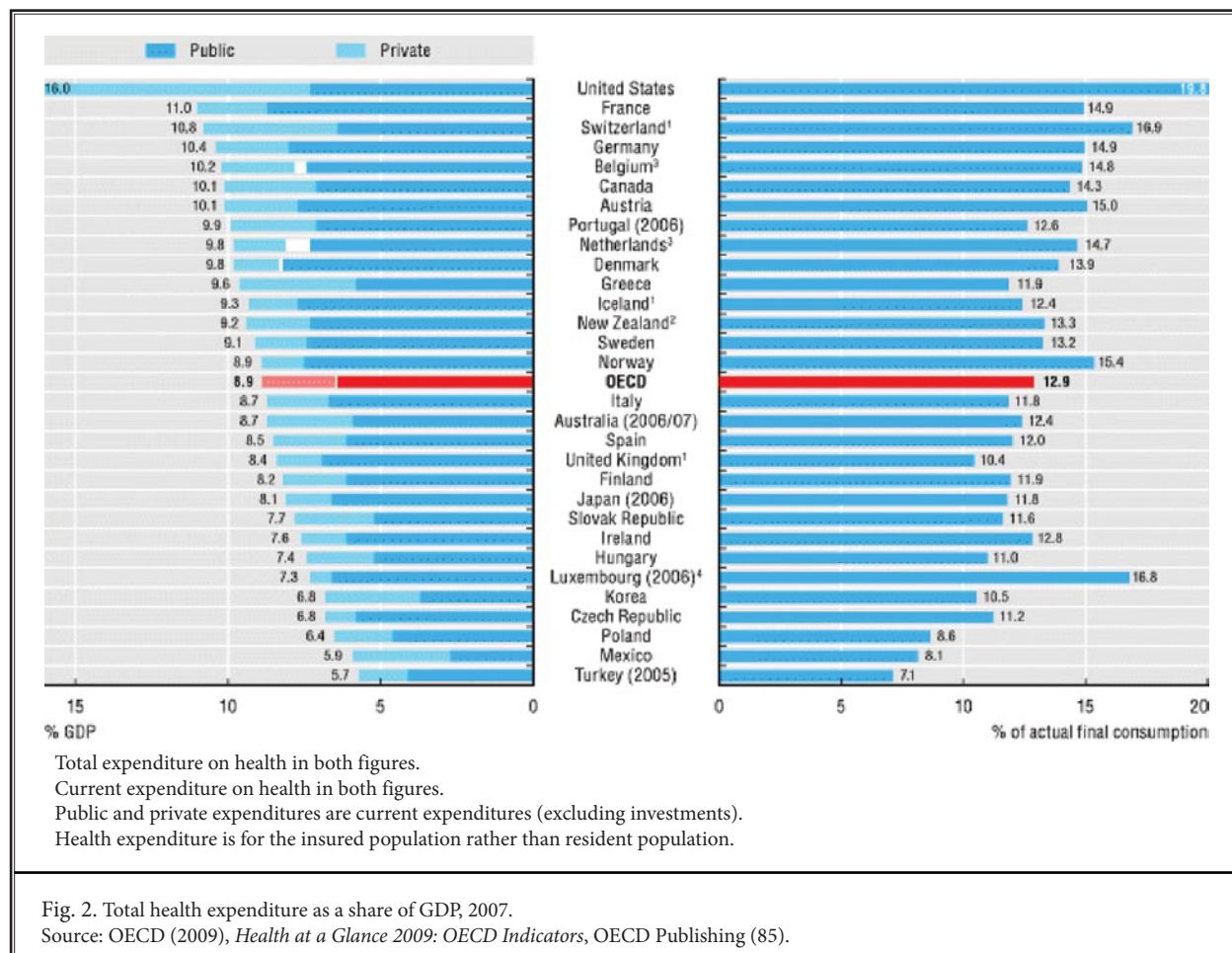


Fig. 2. Total health expenditure as a share of GDP, 2007. Source: OECD (2009), *Health at a Glance 2009: OECD Indicators*, OECD Publishing (85).

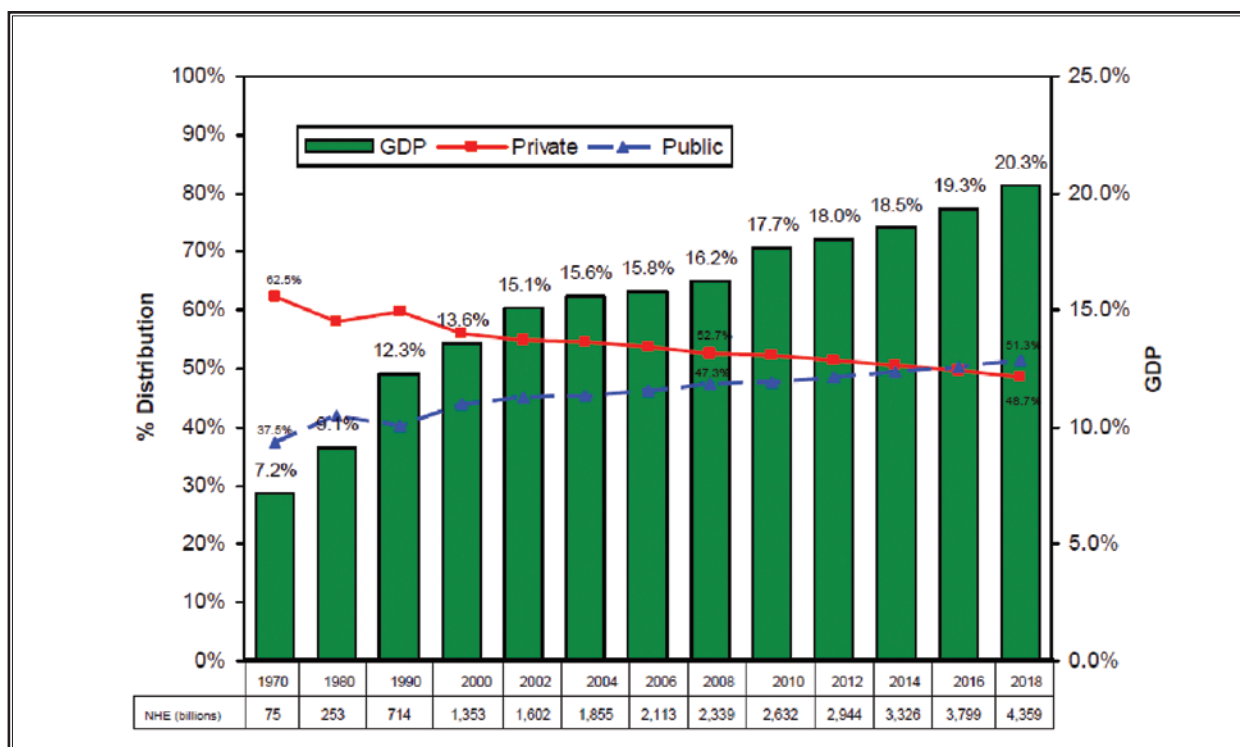


Fig. 3. Increasing health care costs in the United States.

Source: Centers for Medicare and Medicaid Services, Office of the Actuary, National Health Statistics Group, at <http://www.cms.hhs.gov/NationalHealthExpendData> (86) / (Historical data from NHE summary including share of GDP, CY 1960-2008, file nhegdp08.zip; Projected data from “Updated and Extended National Health Expenditure Projections, 2010-2019”, OACT_Memorandum_on_Financial_Impact_of_H_R_3962_11-13-09_.pdf.

in the both public and private sectors representing the same underlying forces, including the development and spread of new and more expensive medical technologies.

However, these estimates have been frequently revised (86). The previous forecasts have shown GDP for 2006 of 16% or higher, for 2007 of 16.2% or higher, and for 2008 16.6% or higher of GDP (Fig. 3). However, actual national health expenditures were less than projected with national health expenses as the percent of GDP of 15.7% in 2005, 15.8% in 2006, 15.9% in 2007, and 16.2% in 2008 (87). In 2008, national health spending reached \$2.3 trillion or \$7,681 per person, an increase of 4.4% from 2007. Even though this is the slowest rate of growth in national aggregate health spending in the national health expenditure accounts (Table 1), the health care portion of GDP grew from 15.9% in 2007 to 16.2% in 2008, despite the downturn in the economy and recession, reflecting the general

pattern that larger increases in health spending share of GDP generally occur during or just after periods of economic recession. Even then, total health care spending was 7.2% of the U.S. economy in 1970.

1.1.1 Public Spending

Over the past 3 decades, federal spending on Medicare and Medicaid has roughly tripled as a share of GDP rising from about 1.8% in 1975 to about 5.7% in 2008 (86,88,89). According to the Congressional Budget Office’s (CBO’s) projections, under the policies in place in 2007, such spending will reach about 12% of GDP by 2050 – but substantial uncertainties surround that estimate (90,91). The CBO report further explained that if cost per enrollee continued growing over the next 4 decades as quickly as they have grown over the past 4 – about 2.5 percentage points faster than per capita GDP – then federal spending on those programs would reach about 17% of the economy. However, if

Table 1. National health expenditures (NHE), aggregate and per capita amounts, and share of gross domestic product (GDP), selected calendar years 1970–2008.

Spending category	1970	1980	1990	2000	2005	2006	2007	2008
NHE, billions	\$74.9	\$253.4	\$714.1	\$1,352.9	\$1,982.5	\$2,112.5	\$2,239.7	\$2,338.7
Health services and supplies	67.1	233.4	666.8	1,264.1	1,851.9	1,975.4	2,089.7	2,181.3
Personal health care (PHC)	62.9	214.8	607.5	1,139.2	1,655.2	1,762.9	1,866.4	1,952.3
Hospital care	27.6	101.0	251.6	416.9	607.5	649.4	687.6	718.4
Professional services	20.6	67.3	216.8	426.8	621.5	658.4	697.5	731.2
Physician and clinical services	14.0	47.1	157.5	288.6	422.4	446.5	472.6	496.2
Other professional services	0.7	3.6	18.2	39.1	55.9	58.4	62.2	65.7
Dental services	4.7	13.3	31.5	62.0	86.3	90.7	96.4	101.2
Other PHC	1.2	3.3	9.6	37.1	56.9	62.7	66.3	68.1
Home health and nursing home care	4.3	20.9	65.2	125.8	168.8	178.1	191.7	203.1
Home health care ^a	0.2	2.4	12.6	30.5	48.1	53.0	59.3	64.7
Nursing home care ^a	4.0	18.5	52.6	95.3	120.7	125.1	132.4	138.4
Retail outlet sales of medical products	10.5	25.7	74.0	169.8	257.4	277.0	289.7	299.6
Prescription drugs	5.5	12.0	40.3	120.6	199.7	217.0	226.8	234.1
Durable medical equipment	1.6	3.8	11.3	19.4	23.8	24.7	25.5	26.6
Other nondurable medical products	3.3	9.8	22.5	29.8	34.0	35.3	37.4	39.0
Program administration and net cost of private health insurance	2.8	12.2	39.3	81.8	140.3	152.0	158.4	159.6
Government public health activities	1.4	6.4	20.0	43.0	56.4	60.6	64.8	69.4
Investment	7.8	19.9	47.3	88.8	130.6	137.1	150.0	157.5
Research ^b	2.0	5.4	12.7	25.6	40.7	41.8	42.5	43.6
Structures and equipment	5.8	14.5	34.7	63.2	90.0	95.3	107.5	113.9
Population (millions)	210.2	230.4	253.8	282.5	295.8	298.8	301.7	304.5
NHE per capita	\$356	\$1,100	\$2,814	\$4,789	\$6,701	\$7,071	\$7,423	\$7,681
GDP, billions of dollars	\$1,038.3	\$2,788.1	\$5,800.5	\$9,951.5	\$12,638.4	\$13,398.9	\$14,077.6	\$14,441.4
NHE as percent of GDP	7.2	9.1	12.3	13.6	15.7	15.8	15.9	16.2
Implicit price deflator for GDP	24.3	47.8	72.2	88.6	100.0	103.3	106.2	108.5
Real GDP, billions chained of dollars	\$4,269.9	\$5,839.0	\$8,033.9	\$11,226.0	\$12,638.4	\$12,976.2	\$13,254.1	\$13,312.2
NHE, billions of 2005 dollars ^c	\$307.8	\$530.6	\$989.1	\$1,526.1	\$1,982.5	\$2,045.9	\$2,108.7	\$2,155.9
PHC deflator ^d	13.3	28.7	58.6	83.0	100.0	103.4	106.9	110.2

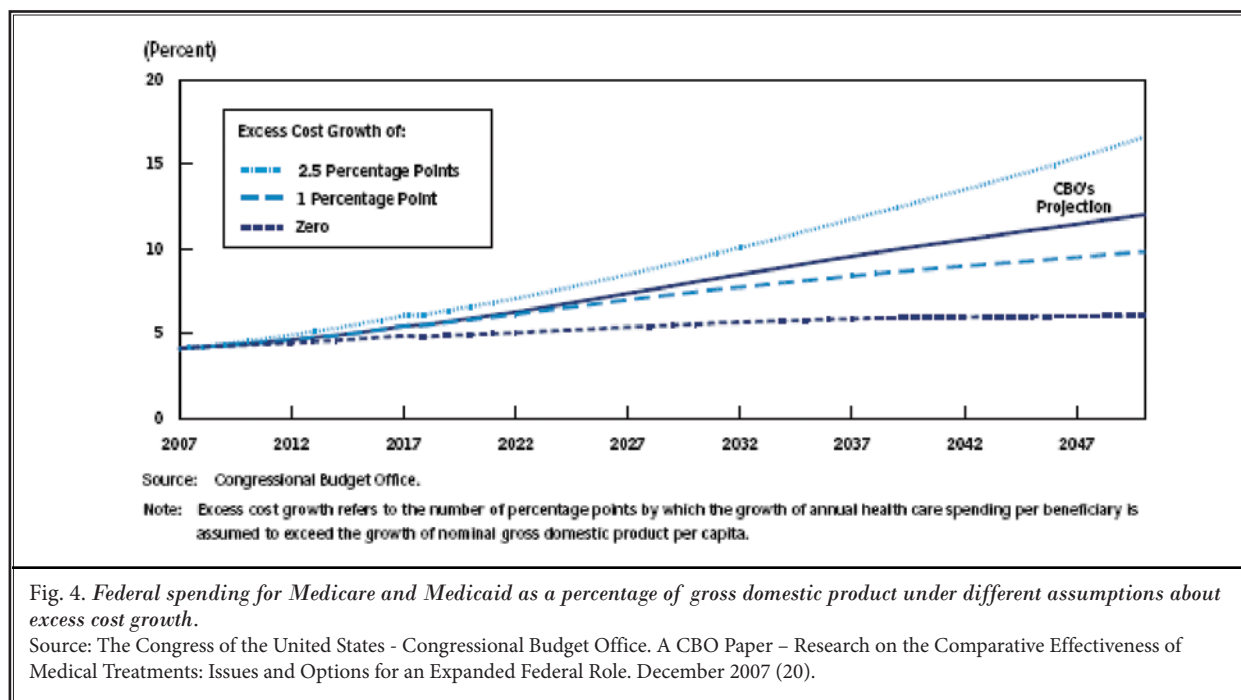
a Freestanding facilities only. Additional services of this type are provided in hospital-based facilities and counted as hospital care.

b Research and development expenditures of drug companies and other manufacturers and providers of medical equipment and supplies are excluded from "research expenditures" but are included in the expenditure class in which the product falls.

c Deflated using the implicit price deflator for GDP (2005 = 100.0).

d PHC implicit price deflator is constructed from the Producer Price Index for hospital care, Nursing Home Input Price Index for nursing home care, and Consumer Price Indices specific to each of the remaining PHC components.

Sources: Centers for Medicare and Medicaid Services, Office of the Actuary, National Health Statistics Group (86); and U.S. Department of Commerce, Bureau of Economic Analysis and Bureau of the Census and Hartman et al. Health spending growth at a historic low in 2008. Health Aff 2010; 29:147-155 (87).



costs per enrollee did not extend the growth of GDP, those federal costs would reach about 6% of GDP in 2050 solely because of demographic changes (Fig. 4). In addition, federal spending has been estimated to account for about one-third of total health care spending. Federal outlays for Medicare and Medicaid alone are projected to nearly double from \$720 billion in 2009 to \$1.4 trillion in 2019. As per the CBO, on a long-term basis, health care costs represent the single greatest challenge to balancing the federal budget (89).

Figure 5 illustrates Medicare expenditures over 10 years from 1999 to 2009 with a cumulative increase of 136% - roughly 8 times the percent rise in Medicare beneficiaries over the same period. However, if these numbers are calculated from 1998 through 2008, the cumulative increase was 124%, roughly 7 times the percent rise in Medicare beneficiaries over the same period with a whole 100% increase with inclusion of either 1998 versus 1999 or 2008 versus 2009.

1.3 Private Spending

Private health care also faces the same challenges as public health care in the United States with rising health care costs (Fig. 3). Consequently, controlling the overall costs over the long-term will be difficult without addressing the forces that are causing private and public health care costs to rise (92).

The premium increases over the last 10 years compared to rise in worker income has escalated substantially (Fig. 6). Over the past decade, from 1999 to 2009, health insurance premiums have risen a cumulative 131% - roughly 4 times the 33% rise in worker incomes over the same period, and even more than the consumer price index (CPI) of 29% (Fig. 6). In addition, the average worker contribution and the average employer contribution also increased similarly from 1999 to 2009 (Fig. 6).

1.3 Health Care Crisis

Rapidly rising health care costs over the decades have prompted the application of business practices to medicine with goals of improving efficiency, restraining expenses, and increasing quality (31,34,93). Concern about escalating costs and the quality of health care delivered in the United States has led up to an increase in focus on pay for performance, value-driven health care, and public reporting of quality and cost information (34).

1.4 Growth of Health Care Services

The need for CER and EBM in clinical practice became a prominent focus during the second half of the 20th century (25). The demonstration of pervasive and persistent unexplained variability in clinical practice,

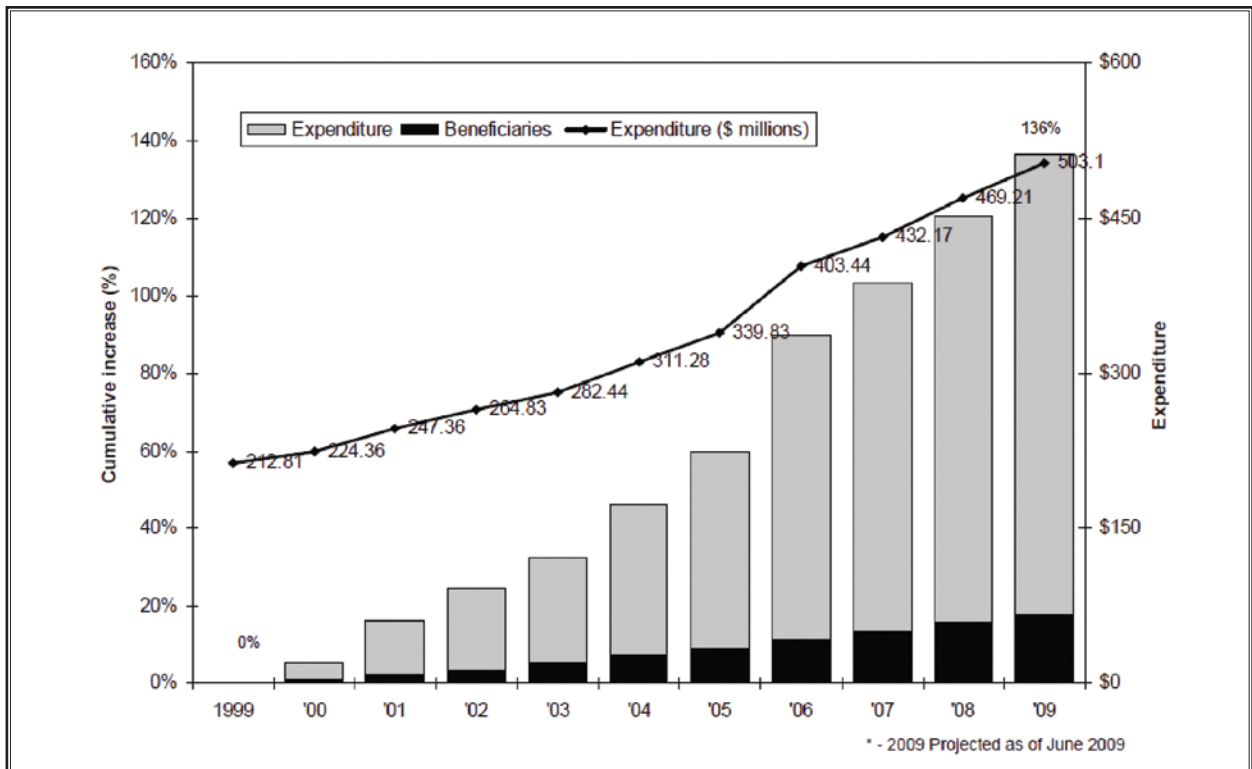


Fig. 5. Medicare expenditure over 10 years (1999-2009*).

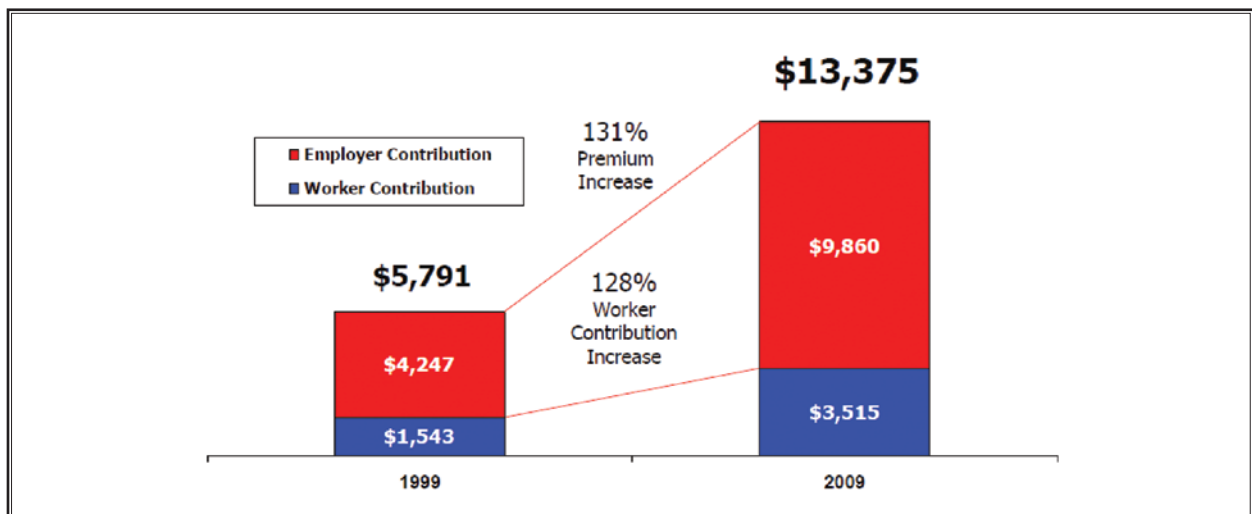


Fig. 6. Average health insurance premiums and worker contributions for family coverage, 1999 – 2009. Source: Kaiser/HRET Survey of Employer-Sponsored Health Benefits, 1999-2009 (92).

and high rates of inappropriate care, combined with increased expenditures, have fueled a steady increase in demand for evidence of clinical effectiveness (26-30). Issues related to chronic pain also take center stage in this debate.

Martin et al (35) evaluated the trends in health care expenditures, utilization, and health status among U.S. adults with spine problems from 1997 to 2006. They reported an increase of 7% per year of national expenditures for spine problems, while population measures of mental health, and work, social, and physical functioning limitations worsened. The largest increase in per-user expenditure was related to prescription medications, increasing 10.2% per year. They showed that the treated prevalence of 13.5% in 2006 is substantially lower than the population prevalence of people with spine problems reported in the national health interview survey of 26%. They illustrated a 139% increase in per-year user expenditures for prescriptions compared to a 40% increase in the number of users reporting prescriptions from 1997 to 2006. Martin et al (37) also reported national expenditures associated with spine problems totaling \$86 billion in 2005, an increase of 65% since 1997.

Deyo et al (38) described overtreatment of chronic back pain including imaging, opioid analgesics, spinal injections, and surgery. The use of magnetic resonance imaging (MRI) increased in the Medicare population by 307% during a recent 12-year interval. In general, spine imaging rates increased in general (39). Further, surgery rates are highest where imaging rates are highest.

Prescription opioid use has been increasing steadily overall; for musculoskeletal conditions in particular (40-45,94-98). Among patients with spinal disorders, the national expenditure panel survey showed a 108% increase in opioid prescriptions from 1997 through 2004 (98). Further, the combination of increasing use and higher drug prices resulted in a 423% inflation-adjusted increase in expenditures. Manchikanti and Singh (40), in a 10-year perspective on the complexities and complications of escalating use, abuse, and non-medical use of opioids, showed an overall increase of 127% in retail sales of opioids from 1997 to 2006 in the United States, with an increase of 1,177% for methadone, 732% for oxycodone, and 479% for fentanyl. Similarly, the increase in therapeutic opioid use in the United States, mg per person, from 1997 to 2006, increased overall 347% with the highest increase for methadone of 1,129% and oxycodone of 899%. Further, emergency department reports of opioid overdose have been in-

creasing in parallel with increasing number of prescriptions, along with diversion.

Deyo et al (38) described that despite no specific concurrent reports of clarified indications or improved efficacy, there was a 220% increase in the rate of lumbar spine fusion surgery from 1990 to 2001 in the United States (46). In addition, the rise accelerated after 1996 when the fusion cage, a new type of surgical implant, was approved. Medicare claims demonstrated a 40% increase in spine surgery rates, a 70% increase in fusion surgery rates, and a 100% increase in use of implants (99). Despite increasing surgical rates and fusions, reoperation rates after initial spine surgery were higher in the late 1990s than earlier in the decade (47).

Similar to lumbar surgery, national trends in spinal fusion for cervical spondylotic myelopathy are rising. The number of patients with cervical spondylotic myelopathy (48) that underwent spinal fusion increased 7-fold from 0.6% to 4.1% per 100,000 people over a period from 1993 to 2002. Further, most spinal fusions were performed in the 45 to 64 year age group.

The Department of Health and Human Services, Office of the Inspector General (HHS-OIG) reported that Medicare paid over \$2 billion in 2006 for interventional pain management procedures (100). This report showed that from 2003 to 2006, the number of Medicare claims for facet joint injections increased by 76%. Manchikanti et al (28-31) in multiple publications have shown an increase in interventional techniques. In the analysis of growth of interventional techniques in managing chronic pain in the Medicare population from 1997 to 2006, overall there was an increase of 137% in patients utilizing interventional pain management services with an increase of 197% in interventional pain management services per 100,000 Medicare beneficiaries (31). The majority of the increases were attributed to exponential growth in the performance of facet joint interventions. Manchikanti et al (31) showed that epidural procedures increased 117%, facet joint interventions increased 543%, discography increased 159%, disc decompression increased 316%, spinal cord stimulation increased 518%, and all types of other nerve blocks increased 63%, whereas, intrathecal infusion pumps increased only 29% per 100,000 Medicare beneficiaries from 1997 to 2006. Vertebral augmentation procedures increased 218% from 2002 to 2006. Overall, per 100,000 Medicare beneficiaries, the increases were 197%. Friedly et al (33) illustrated a 271% increase in lumbar epidural steroid injections from 1994 to 2001 and a 231% increase in facet joint injections.

The RAND health investigation of 10 physicians' services (32) examined the underlying growth in 10 Selected High Growth Service Categories that saw significant increase (40%) for overall growth in allowed charges among Medicare beneficiaries between 2000 and 2006. This evaluation included lumbar transforaminal epidural injection, as well as lumbar facet joint nerve block. The allowed charges per beneficiary increased 731% (122% per year) from 2000 to 2006.

1.5 Geographic Variation

Another important effect of limited evidence is geographic variation. It has been suggested that substantial evidence on the variations in medical care from area to area around the country indicates that as much as 30% of spending reflects medical care of uncertain or questionable value (71). Investigators at Dartmouth have documented significant geographic variations in the intensity of services for colorectal cancer, hip fracture, acute myocardial infarction, coronary artery bypass grafting, hysterectomy, lumbar surgery, and end-of-life care (49-53,101,102). Further, intensity of discretion of services such as lumbar surgery, hysterectomy, and bypass surgery can vary by as much as a factor of 20 depending on the location (72,103). It has been

shown that, in Idaho Falls, Idaho, 4.6 lumbar fusions were reported per 1,000 Medicare enrollees annually compared to 0.2 in Bangor, Maine, with no difference in the outcomes (72,103). Thus, it has been concluded that many of the medical treatments in common use, as well as many emerging therapies, are not backed by strong empirical evidence – leading to the conclusion that less than 50% of treatments delivered today are supported by evidence (71,73). These differences and lack of evidence is most distinct in end-of-life care. Such patients spend nearly 20 days in the hospital over those last 6 months, on average, in highest use areas, compared with an average of about 6 hospital days in the lowest use areas. Further, the average number of visits to physicians in that period is as high as 15 in some of the highest-use regions, and as low as 16 in some of the lowest-use regions (26).

The observed variations in the use of services correspond to substantial differences in Medicare spending per enrollee in different parts of the country (Fig. 7) (20,54). In 2006, the average costs ranged from about \$5,542 in the areas with the lowest spending, compared to \$16,351 in the areas with the highest spending. Researchers at Dartmouth described that differences in illness rates account for less than 30% of the variation

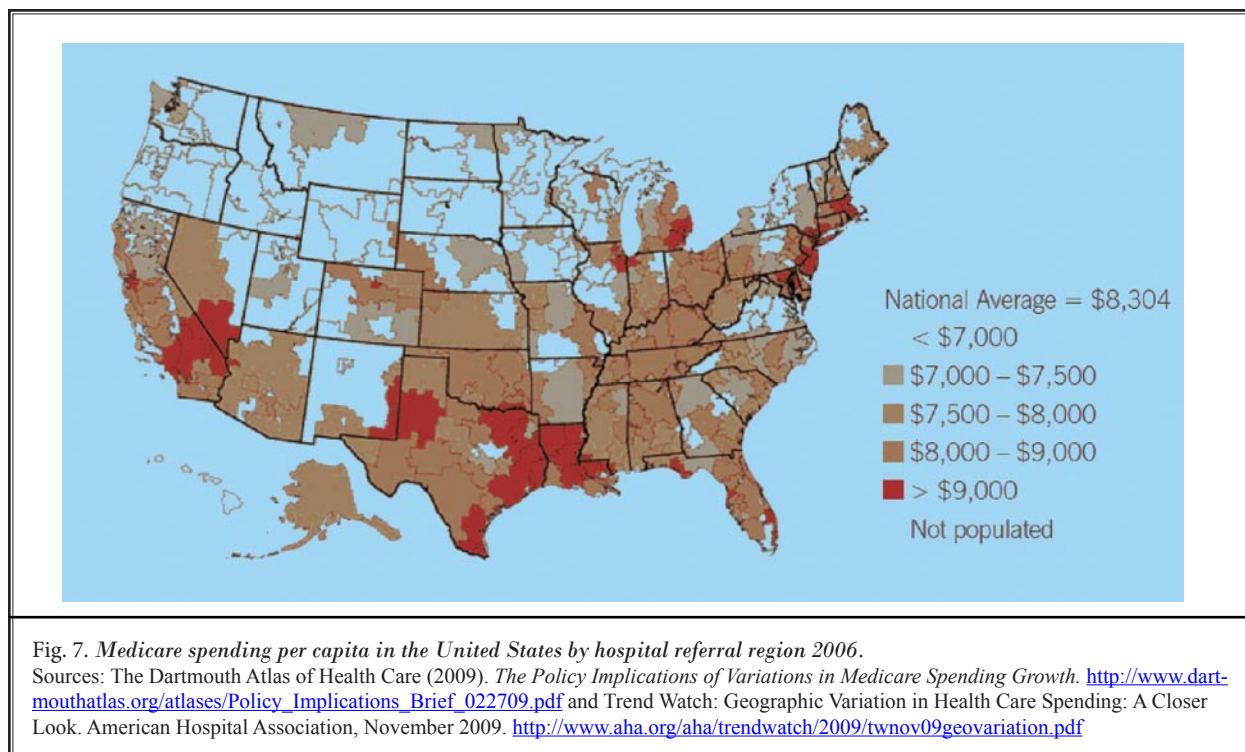
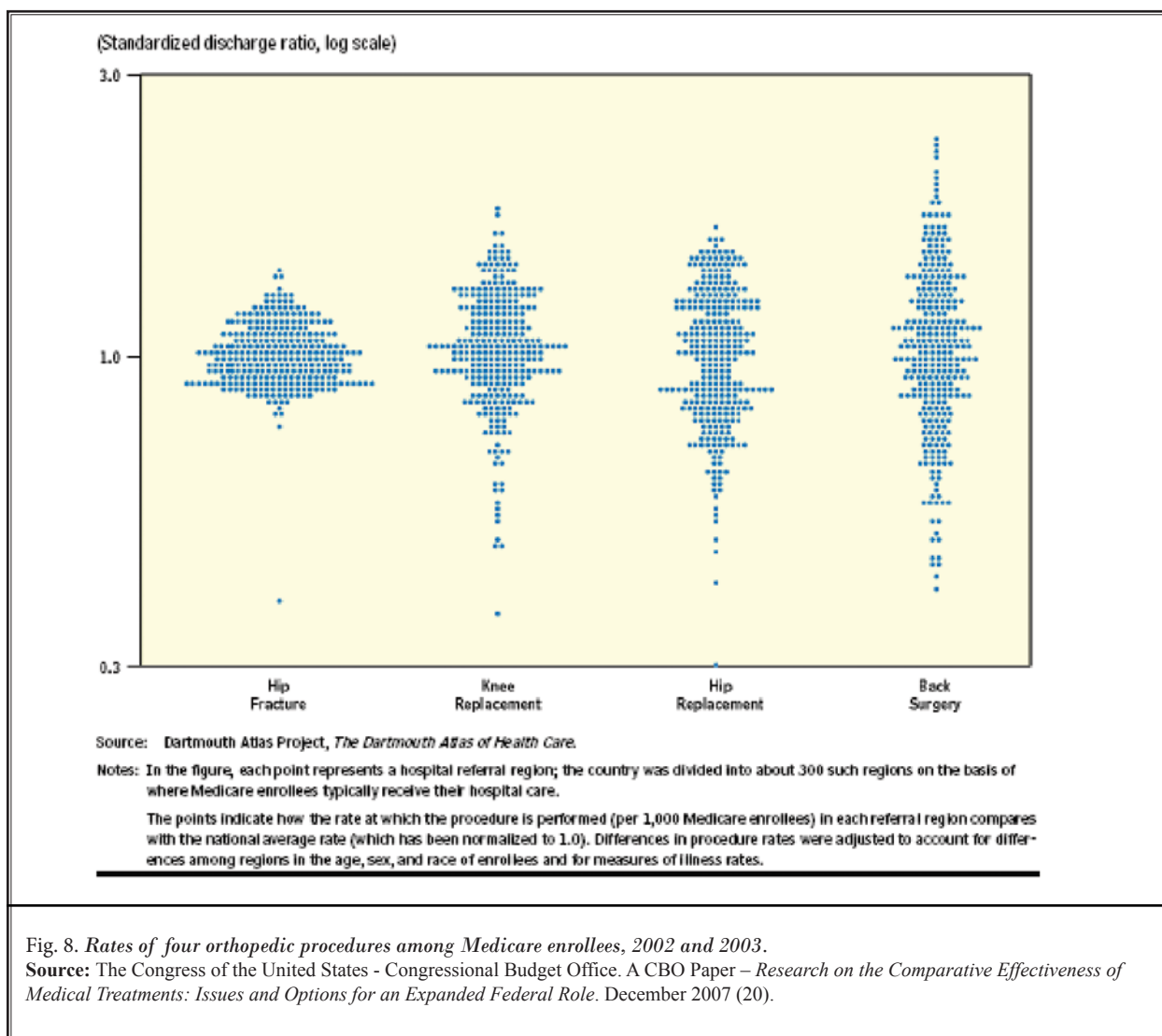


Fig. 7. Medicare spending per capita in the United States by hospital referral region 2006.

Sources: The Dartmouth Atlas of Health Care (2009). *The Policy Implications of Variations in Medicare Spending Growth*. http://www.dartmouthatlas.org/atlas/Policy_Implications_Brief_022709.pdf and Trend Watch: Geographic Variation in Health Care Spending: A Closer Look. American Hospital Association, November 2009. <http://www.aha.org/aha/trendwatch/2009/twnov09geovariation.pdf>



in spending among areas, and differences in prices can explain another 10% - indicating that more than 60% of variation is due to other factors (49). However, other studies have shown that a larger share of variation in spending can be accounted for by differences in health status and demographic factors, but even so, the remaining differences are substantial in dollar terms (55).

It has been stated that there is evidence that the degree of geographic variation in treatment patterns is greater when less consensus exists within the medical community about the best treatment to use. As an example of good evidence, it is stated that patients who have fractured a hip need to be hospitalized; there is

relatively little variation in admission rates for Medicare beneficiaries with that diagnosis. However, for hip replacements and for knee replacements, more discretion is involved and the surgery rates vary more widely (Fig. 8). Furthermore, the variation in the rates of back surgery and other treatments, whose benefits have been subject of substantial questions from within and outside, show the major variations.

Significant geographic variations have been demonstrated with all types of spinal interventions, including interventional techniques. Spine imaging rates vary dramatically across geographic regions, and surgery rates are highest where imaging rates are highest (104). Further, when judged against guidelines, one-third to

two-thirds of spinal computed tomography (CT) and MRI may be inappropriate (56-58,105). Significant variations have been reported with lumbar spine surgery. Further, higher spine surgery rates are sometimes associated with worse outcomes. In the state of Maine, the best surgical outcomes occurred where surgery rates were lowest; the worst results occurred in areas where rates were highest (58,99).

Deyo and Mirza also (106) evaluated the trends and variations in the use of spine surgery. They showed that spine surgery rates in the United States increased 55% in the 1980s. Among Medicare patients there was a 6-fold variation in spine surgery rates among United States cities, and a 10-fold variation in spine fusion rates. The most rapid increase was for spinal fusion, which tripled during the 1990s and accounted for an increase in proportion of all spine procedures. Further, some increases coincided with the introduction of new surgical implants. However, all the issues may not be scientifically explained just based on geographic variation.

The rates also varied for all degenerative spinal disorders along with specific increases for cervical spine (107-112). In contrast, even though numerous variations have been observed with surgery of the hip and knee, variations in back surgery were the highest (113).

Friedly et al (59) showed significant geographic variation in epidural steroid injections among Medicare patients in the United States. They reported that in 2001, there was a 7.7-fold difference between the state with the lowest rate, Hawaii (5.2 per 1,000), and the state with the highest rate, Alabama (39.9 per 1,000). The variation among health referral regions, which are small in size, was even greater, with an 18-fold difference from 5.6 per 1,000 in Honolulu, Hawaii, to 103.6 per 1,000 in Palm Springs, California. Higher statewide rates of epidural steroid injections were associated with significantly higher rates of lumbar surgery.

Manchikanti et al (31) evaluated specialty characteristics as well as characteristics by each state. The average increase of services from 1997 to 2006 was 197% for the United States, an annual percent increase of 19.7%. There were several states with increases greater than 100% average (i.e., greater than 297%). In contrast, there were also decreases seen in 2 states with California (37% increase), and Idaho with an 81% increase, with some states showing increases of 6% to 95%. Overall, there was 13.9-fold difference between the state with the lowest rate, California (37%), and the state with the highest rate, Connecticut (514%), from 1997 to 2006. Florida had an 11.6-fold increase com-

pared to California (431% versus 37% increase). The average difference for the United States for services per 100,000 beneficiaries was 4.8-fold with the lowest rate of services in Hawaii. Fifteen states had above average increases. Florida presented with 27,979 services per 100,000 Medicare beneficiaries compared to Hawaii with 22,101 services per 100,000 Medicare beneficiaries, a 12.7-fold geographic difference.

The implications of the observed variations in treatments and spending depend importantly on their relationship to health outcomes. If life expectancy and other measures were better in the areas with higher spending, that result would imply that increased spending the low-cost areas would yield health benefits. A study examined differences in hospital spending in Florida and found that areas with higher spending had lower mortality rates among Medicare patients who were treated in the emergency room for heart attack (60). However, another study found that higher-spending regions did not, on average, have lower mortality rates than the lower-spending regions, even after adjustments to control for differing illness rates among patients and regions (61). Further, this study also found that higher spending did not slow the rate at which the elderly developed functional limitations. Even though more research is needed about the impact that differences in spending have on patients' morbidity and quality of life, perhaps using more sophisticated and expensive health care measures may or may not be responsible for variations. In addition, suggestion of these findings with spending in high-cost areas could be reduced without adverse effects on the overall health of residents in those areas, is not based on EBM or CER.

1.6 Why Variations?

Why is there so much practice variation? While it may be easy to dismiss it as "the art of medicine," the main causes are uncertainty and lost translation (17). Uncertainty because for many common clinical scenarios, definitive evidence demonstrating that one approach is better than another does not exist. This extends not only for specialties such as interventional pain management, but also to a well established specialty such as cardiology. For example, it is not known if a patient with new onset of chest pain will have a different outcome if referred for one type of non-invasive test as compared to another (61). Even though cholesterol drugs are one of the most commonly used drugs in the United States, it is not certain which anti-cholesterol

drugs are most effective or cost effective, and the problems of hypercholesteremia are unresolved (62,76).

A review of practice guidelines developed by the American College of Cardiology (ACC) and the American Heart Association (AHA) found that relatively few recommendations were based on high quality evidence (114). Another study also revealed that most guidelines for treating lung cancer were not based on adequate evidence (115). It has been stated that a major reason for the gap in the evidence is because of the lack of comparative effectiveness and very low allocation of funds (currently less than 0.1%) in assessing the comparative effectiveness of available interventions (116,117). Further, the absence of timely and relevant evidence appears to be a major issue.

While an evidence gap is an attractive argument, improper use of evidence with inappropriate synthesis, bias, and even outdated evidence, is a major problem (118-151).

Some explanation may be provided for geographic variations in performance of interventional techniques such as lack of appropriate regulations, lack of training requirements, lack of fluoroscopy, lack of utilization of appropriate indications, and medical necessity. Finally, economic incentives have been fueling the costs in some regions, such as Florida. All the causes may not be applicable. Further, gaming of the studies with inappropriate interpretation may also lead to misrepresentation of geographic variations and effectiveness of a modality. For example, it has been shown that patients who receive high dose opioids more frequently undergo epidural injections as well as surgical interventions; however, the factors such as functional disability and intensity of pain and structural abnormalities have not been taken into consideration in such conclusions (59). The same applies to functional disability on opioids or patients undergoing injection therapy (63,152-154). Further, it has been shown that patients who have undergone epidural steroid injections are more likely to undergo surgical decompression for spinal stenosis. Epidural steroids have been blamed for the increase in surgery. However, the reasons in the first place to undergo epidural steroid injections is the level of symptomatic stenosis and functional disability (59). To be appropriate, one should look at the proportion of patients undergoing epidural steroids who were able to avoid surgical intervention and the resulting cost effectiveness.

Health policy experts cite practice variation as symptomatic of uncertainty, waste, inefficiency, and poor performance of the health care enterprise (17,64,155).

2.0 EVOLUTION OF CER

While press attention has gone to the American Recovery and Reinvestment Act (ARRA) of 2009, and to President Obama for comparative effectiveness, it started long before 2009, with passage of the Medicare Modernization Act (MMA) in 2003 (19,156). In fact, the U.S. government has a rather long, but somewhat checkered history of involvement in CER and related efforts.

Historically, the intellectual roots of effectiveness research can be traced back to the mid 18th century Scotland and the "arithmetical medicine" practiced by the graduates of Edinburgh Medical School (1). James Lind at this medical school undertook a controlled trial of 6 separate treatments for scurvy (2). During the 1830s, Pierre Louis developed the *méthode numérique* in Paris, whereby he demonstrated that phlebotomy did not actually improve the survival rates of patients suffering from pneumonia (1). At the beginning of the 20th century, Ernest Codman, an American physician, founded what is today known as "outcomes management" in patient care (3). He published his results in a book, *A Study in Hospital Efficiency* (3). Of 337 patients discharged from the hospital between 1911 and 1916, Codman recorded and publicized 123 errors.

In England, the 1930s saw the development of health services research in a world increasingly obsessed with egalitarian uniformity. J.A. Glover found a tenfold variation in tonsillectomy (65). Subsequently, following several decades of socialized health care in the United Kingdom, the 1970s and 1980s witnessed the release of a range of studies that highlighted the wide geographical variations in general medical admissions, including operations such as appendectomy, caesarean section, cholecystectomy, hysterectomy, tonsillectomy, and prostatectomy (66). Such variations not only demonstrated the inequities of the National Health Services (NHS), but also raised questions about the probity and cost effectiveness of many of its treatments.

Following the publication of Archie Cochrane's *Effectiveness and Efficiency: Random Reflections on Health Services* (67) in the United States, researchers demonstrated large variations in the rates of prostatectomy for patients with benign prostatic hypertrophy (68). Consequently, the opinions were drawn that such variation meant either under-provision in some places and/or over-provision and possibly ineffective treatment in other regions (157).

2.1 CER in the United States

In the United States, federal efforts date at least to the late 1970s and the short-lived National Center for Healthcare Technology. It was established in 1978 as part of the Department of Health, Education, and Welfare (DHEW) and was given a broad mandate to conduct and promote research on health care technology. It included an Advisory Board appointed by the secretary to assist in setting research priorities. The Center sponsored or co-sponsored major evaluations of coronary artery bypass graft surgery, dental radiology, and caesarean delivery and made about 75 recommendations to the Medicare program about coverage (20). The Center ceased operations at the end of 1981, reflecting changes in priorities for the new administration and the Congress, as well as opposition from some provider and industry groups (158).

In 1972, the Office of Technology Assessment (OTA) was created as an advisory agency to Congress, covering a broad set of issues, including health care. Most of the focus on evaluation of technologies now would be called CER. The OTA produced an extensive review and analysis of the issues involved in options for improving evidence about the clinical effectiveness and cost-effectiveness of medical treatments (159). For a variety of reasons – however, having little to do with its health care study specifically but instead reflecting broader questions about the agency's role – the OTA was eliminated in 1995.

In 1989, the Agency for Healthcare Policy and Research (AHCPR) was created as an arm of the Department of Health and Human Services (DHHS) (160). AHCPR has undertaken a number of initiatives, including creation of the National Guideline Clearinghouse (NGC) designed to summarize the available medical evidence on the appropriate treatments for various conditions (160). They produced 15 guidelines at a cost of \$750 million. In the mid 1990s, controversies arose after an agency-sponsored research team concluded that there was insufficient evidence to support certain spinal surgeries, and on the basis of that, the agency issued practice guidelines for the treatment of back pain (81,161-163). Strong opposition from spine surgeons, along with broader questions about the value of the research that the agency had funded and other factors, led to pressure to eliminate the agency (163).

Ultimately, AHCPR was retained, but its funding for fiscal year 1996 was reduced from prior levels. It was renamed the Agency for Healthcare Research and Quality (AHRQ). Since then, its overall budget has gen-

erally been maintained, at least in nominal terms, or increased (160).

In 2003, the landmark MMA authorized AHRQ to spend up to \$50 million in 2004 and additional amounts in future years to conduct and support research with a focus on "outcomes, comparative clinical effectiveness, and appropriateness of healthcare items and services" for Medicare and Medicaid enrollees (156,164,165). The actual funding appropriated for that initiative has been \$15 million per year. Using that funding, AHRQ has established an "effective healthcare" program consisting of 3 main functions: reviewing and synthesizing existing evidence (using its evidence-based practice centers); generating new information using a set of approved research centers (such as the HMO research network) that have access to data from medical claims and electronic medical records; and publishing findings and formats that are geared to the differing needs of clinicians, patients, and policy-makers (166).

Other federal agencies also engaged in various activities related to CER – efforts that received less attention than AHRQ's activities, but that are probably larger in dollar terms. The Department of Veterans Affairs (VA) has a very substantial research program that reviews evidence from the medical records of its patients, focusing particularly on the clinical effectiveness of treatments. The department also sponsors evidence reviews through the technology assessment program and helps fund clinical trials – including the study comparing strengths to drug therapy. Over the years, the NIH has sponsored a number of trials that compare treatments directly.

The Centers for Medicare and Medicaid Services (CMS) also has helped to sponsor a limited amount of research on comparative effectiveness (i.e., lung volume, reduction surgery). CMS generally considers only whether devices and procedures are clinically effective in making payment decisions. However, it has sponsored some studies comparing the effectiveness of different treatments but has done so largely to determine whether to establish separate payment rates for similar treatments (20). CMS has sponsored a trial with NIH that may eventually compare the effects of daily dialysis for kidney patients with the conventional treatment of dialysis 3 times per week. However, this may increase the cost rather than reduce it if daily dialysis proves more effective for certain patients. CMS could modify its payment policy to cover the additional costs of more frequent treatments for those patients.

Overall, it has been estimated that the federal government has spent \$1.5 billion in 2005 on health services research. This broader category includes some of the work on comparative effectiveness, but also encompasses many other types of studies (117). Further, it also has been stated that aggregate figures may not include all federal funding for comparative trials or other efforts that are outside the traditional scope of health services research.

2.1.1 American Recovery and Reimbursement Act

On February 17, 2009, President Obama signed into law the ARRA. Of \$787 billion that was appropriated, \$150 billion was allotted for medical issues, being touted as a down payment on health care reform (167,168). Among the most controversial provisions of the medical spending was the allotment of \$1.1 billion for CER. Further, multiple bills have been introduced in the 100th U.S. Congress that directly address CER (17). Senator Max Baucus (D-MT) introduced the Comparative Effectiveness Research Act of 2008 which proposed to establish a private, non-profit corporation called the Health Care Comparative Effectiveness Research Institute, which never became law (169). The institute would be governed by a board with representatives from multiple sectors. It would be charged with identifying national priorities for CER, and would be allowed to enter into contracts with different entities for conducting research.

2.3 CER Internationally

CER has been the basis of decision for health care in many other countries (1). According to the International Network of Agencies for Health Technology Assessments (INAHTA) (170), many industrialized countries have bodies that are charged with health technology assessments (HTAs) or comparative effectiveness studies. However, the evolution of these bodies and their responsibilities at the national decision-making level has been far from uniform. While some of these bodies have an advisory role and make reimbursement or pricing recommendations to a national or regional governing body, others have a more explicit regulatory role. They are accountable to government ministries and are responsible for listing and pricing medicines and devices. This is the case in France, Germany, and the United Kingdom.

2.2.1 CER in United Kingdom

The NHS was established in 1948 in the United Kingdom. It is a single payor health care system, directly ad-

ministered by the British government, funded through taxation, and provided mainly by public sector institutions (4,171). In 1999, the government established the NICE (172-174). At its heart is the center for health technology evaluation that issues formal guidance on the use of new and existing medicines based on rigid and proscriptive "economic" and clinical formulas. The NHS is obliged to adhere to NICE's pronouncements.

2.2.3 CER in Denmark

The Danish health care system is completely state-funded, with public provision of hospital beds representing more than 90% of the hospital sector. Denmark's national HTA system was explicitly established on the basis of its making prioritized resource-allocation decisions carried out by the unit known as the Danish Centre for Evaluation and Health Technology Assessment (DACEHTA). It operates within the framework of the National Board of Health (NBH), itself a part of the Danish Ministry of Health (175).

2.2.3 CER in France

In France, in 2005, a centralized High Health Authority was established. It is designed to stipulate the benefits of medicines and determines their price-reimbursement levels. As such, it is set to raise the focus on cost-containment and bring its decision-making under closer state control (173,174).

2.2.4 CER in Germany

In Germany, as in France, health care is financed primarily by social insurance and provided by a mixture of public and private providers. However, only 10% of Germans opt for full private medical insurance (172-174,176-178). In 1990, the Office of Technology Assessment at the German Parliament (TAB) was established, and in 2004, the government set up the Institute for Quality and Economic Efficiency in the Healthcare Sector (IQWiG). Tasked with the central goal of efficiency, the IQWiG investigates and stipulates which therapeutic and diagnostic services are appropriate.

2.2.5 Overall European Perspective

The European community has promoted priority setting, effectiveness assessments, information sharing, and the dissemination of results since 1994 (173). The European community established the European Network for Health Technology Assessment (EUnetHTA) in 2006 to promote better coordination of national efforts (174). Essentially, this Europe-wide initiative serves

as an umbrella effort to make certain that there is no duplication of efforts and to bring up standards across individual countries and agencies.

2.2.6 International Perspective

Systematic, detailed information on the operations of most national clinical effectiveness programs is limited, and studies assessing and comparing the impacts of these programs are even more limited (176). Further, the documentation and evaluation of national programs assessing clinical effectiveness that are consistent, transparent, and evidence-based are not uniform (173-178). However, even the IOM Committee has not undertaken an in-depth study of international models for developing knowledge about clinical effectiveness (179).

As shown in Table 2, effectiveness review programs in Australia, Canada, Denmark, France, Germany, and the United Kingdom assess a broad range of clinical services, including drugs, devices, tests, imaging procedures, preventive services, and surgical procedures (9,21,180-199). The programs in Australia, Canada, Germany, and the United Kingdom assess both clinical effectiveness and cost-effectiveness.

3.0 WHAT IS COMPARATIVE EFFECTIVENESS RESEARCH?

3.1 Definition of CER

An agreed upon definition of CER is an essential first step for setting priorities and developing a sustainable national CER (4). It informs the public of the focus of this research and its importance in their lives, and it informs investigators of the characteristics of the research to be supported by CER funds. Further, it provides a basis for judging research proposals to perform CER and for evaluating the impact of that research and the success of a national CER program. In formulating its definition, the Committee on Comparative Effectiveness Prioritization, IOM (4), drew upon definitions by several government agencies and other IOM committees:

Comparative effectiveness research (CER) is the generation and synthesis of evidence that compares the benefits and harms of alternate methods to prevent, diagnose, treat, and monitor a clinical condition or to improve the delivery of care. The purpose of CER is to assist consumers, clinicians, purchasers, and policy-makers to make informed decisions that will improve health care at both the individual and population levels.

Table 2. Focus of selected national efforts to identify effective health care services.

Country	Drugs	Devices ^a	Preventive Services	Surgical Procedures ^b
United Kingdom	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
Australia	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Canada	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Denmark	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
France	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Germany	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Scotland	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
England and Wales	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

^aIncludes diagnostic and therapeutic devices (e.g., ultrasound machines, stents, and inhaler devices).

^bIncludes the assessment of operating techniques, the use of surgical equipment for a specific procedure, and comparative effectiveness of surgical procedures.

Data adapted from: Australian Safety & Efficacy Register of New Interventional Procedures-Surgical (2005) (180); CADTH (2006) (181); Canadian Task Force on Preventive Health Care (2005) (182); Department on Health and Ageing (2006) (183); Haute Autorité de Santé (2007) (184); Institute for Quality and Efficiency in Health Care (2007) (185); National Board of Health (2007) (186); NICE (2007) (187); SIGN (2007) (188); National Health and Medical Research Council (2006) (189).

Source: Eden J et al. *Knowing What Works in Health Care: A Roadmap for the Nation*. National Academies Press, Washington, DC, 2008 (9).

3.3 Goals of CER

The goals of comparative effectiveness are not only to compare research of alternative approaches to prevent, diagnose, treat, and monitor clinical conditions, but also to support patient-centered care that will produce superior patient outcomes. CER is expected to produce important clinical innovations more rapidly and uniformly and at the same time holding the health care costs to a lower level to ensure that the highest value is obtained for every health care dollar (190). To achieve these goals, comparative effectiveness will require a clear, unbiased, non-bureaucratic strategy to produce, disseminate, and apply the research to help all stakeholders, in order to improve the quality and value of clinical practice.

3.3 Functions of CER

CER combines 2 very separate elements; first synthesizing existing evidence to inform decision-making; second, to generate new evidence to address important evidence gaps. Synthesizing evidence can be accomplished through systematic reviews and meta-analysis; however, the generation of new evidence requires the commissioning of prospective clinical trials or the conduct of new analysis of patient outcomes from data available in insurance claims systems, electronic medical records, or clinical registries. These databases, however, seldom specify the rationale for medical decisions. Prospective registries and cohort studies are undertaken to understand the natural progression of disease and factors that influence clinical outcomes. These observational research methods have many advantages, but cannot escape a key limitation: characteristics of the patient drive real-life clinical decisions, leading to uncertainty about whether they, or the intervention itself, cause the observed outcomes. Therefore, researchers often turn to RCTs, in which patients are randomly assigned to different interventions eliminating much of the uncertainty that plagues the interpretation of the observational research. Over time, differences in response to these interventions reveal which work best and identify factors that might predict the benefits or harms of an intervention. Even though RCTs are considered the gold standard of evidence, they have shortcomings. Researchers must choose among these methodologies and must inform the public about their shortcomings.

When combined effectively, the elements of CER represent a powerful tool for improving the evidence base and informing decisions made by patients, clinicians, and policy makers (190). However, the under-

standing of the functions of CER has been lacking due to expected bias, uncertainty, and political tensions surrounding the question of how comparative effectiveness information will be used. Consequently, rhetoric has been projected on both sides with only 2 options for the function of a comparative effectiveness program: one that focuses narrowly on providing information to patients and clinicians, or one that largely serves the interests of payors by making or recommending coverage decisions. However, comparative effectiveness evidence should achieve a broader function when it provides patients and clinicians with more appropriate information while also providing payors evidence framed specifically to support value based coverage and payment policies. Thus, as Pearson (190) describes, the function of CER is to guide practice and policy from better evidence to better care.

4.0 COMPARATIVE EFFECTIVENESS RESEARCH VS. EVIDENCE-BASED MEDICINE

The CER and the EBM, though similar, are not synonymous. Thus, it is not only the definitions, but also the methodologies and application. While they vary, both in principle are about providing high quality evidence.

The EBM has been defined as a conscientious, explicit, and judicious use of current best evidence in making decisions about the care of individual patients (5). In contrast, comparative effectiveness is the generation and synthesis of evidence that compares the benefits and harms of alternative methods to improve the delivery of care, RCTs, systematic reviews, meta-analysis, observational studies, and all types of evidence. Further, EBM systematic reviews utilize explicit methodology of clearly formulated questions and methods to identify, select, and critically appraise the relevant research and then collect and analyze the data from the studies that are included in the review (5-7,10-15). In contrast, a CER is a unique type of systematic review, which synthesizes the available scientific evidence on a specific topic. The CER expands 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 (4,73). Consequently, it is stated that in doing so, CERs more closely parallel the decisions facing clinicians, patients, and policy-makers, who must choose among a variety of alternatives in making diagnostic, treatment, and health care delivery decisions (73,191-193).

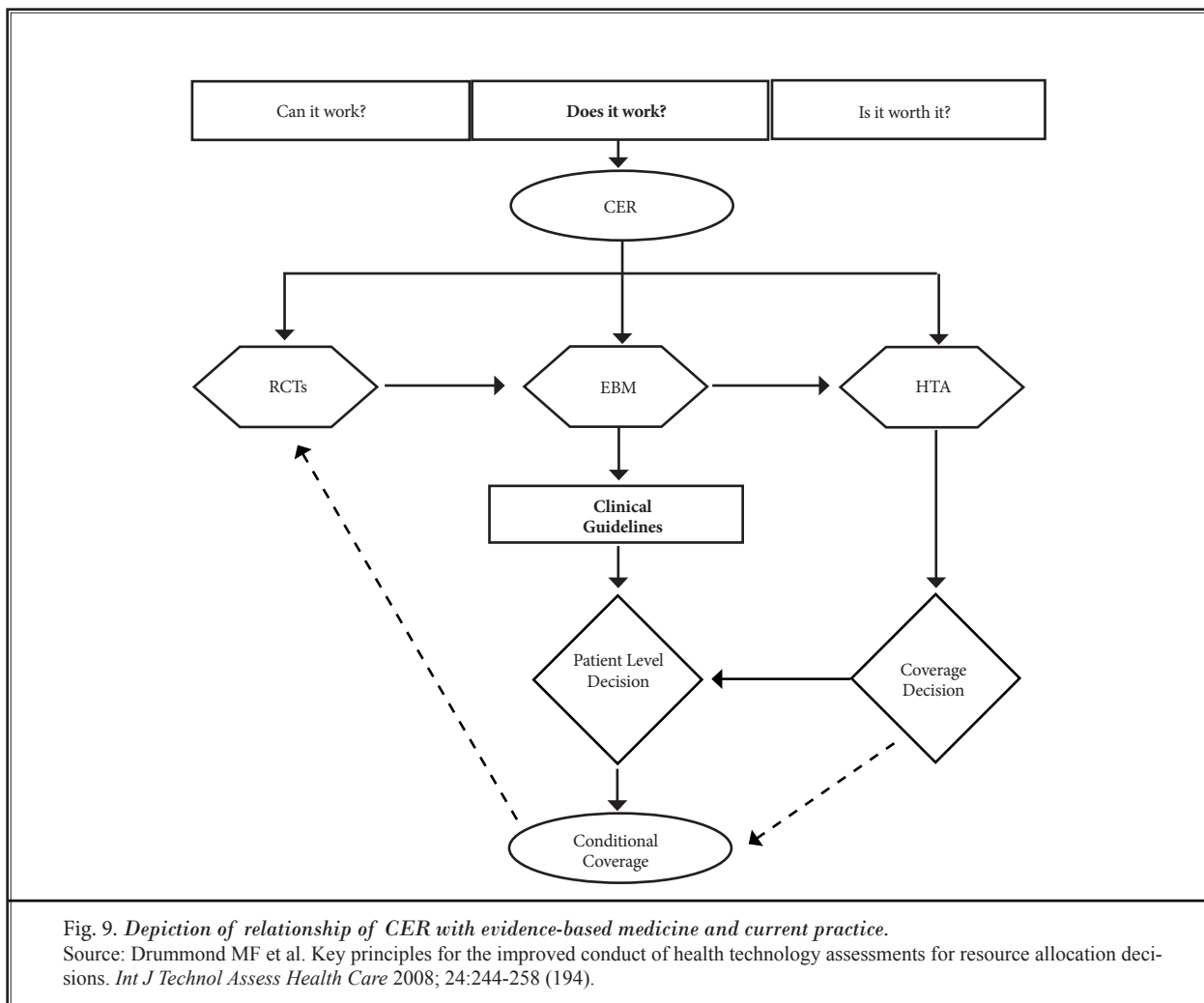
Drummond et al (194) has described the role of CER in current practice in relation to EBM (Fig. 9).

5.0 THE FEDERAL COORDINATING COUNCIL

5.1 The Structure of the Council

The ARRA established the Federal Coordinating Council for CER to foster optimal coordination of CER conducted or supported by the federal government. The council consists of 15 members, all of whom must be government employees and at least half of whom must have clinical experience. Ezekiel Emanuel, the brother of White House Chief of Staff Rahm Emanuel, is a member of the council (18). In his book *Healthcare Guaranteed*, he wrote, "The Institute for Technology and Outcomes Assessment will provide information on

effectiveness and cost to eliminate tests or treatments of marginal or no value." This council bears a striking resemblance to the agency that assesses comparative effectiveness in Great Britain - NICE. It is worth stressing that NICE analyzes both clinical effectiveness and cost effectiveness. In addition, in his book *Critical*, former Senate Majority Leader Tom Daschle extols a health care system based on value (195). It has been interpreted that this is simply another way of saying that the government should be able to determine health care coverage based on its rigid criteria leaving little room for the professional judgement of the physician or the unique biology of the patient. Daschle wrote, "Doctors . . . hospitals and other health care providers will have to adjust to an underlying value-oriented system. In too many cases, they are providing care that does



not reflect the latest science. Doctors will have to learn to operate less like solo practitioners and more like team members. In return, they will enjoy the benefits of working in a simpler seamless system that requires excellent performance.”

5.3 Recommendations

On June 30, 2009, the council released a report to President Barack Obama and Congress on its recommendations for CER funding priorities for the Office of the Secretary (196,197). The council established explicit threshold and prioritization criteria to guide recommendations for funding priorities as shown in Table 3. The council also developed a strategic framework for categorizing current CER activity, identifying gaps, and informing the recommendations for priorities.

The council summarized that the expansion of CER or patient-centered outcomes research has at least 3 major implications. First, the results of such research will better inform a broad area of health care decision. Second, the ARRA provision for CER represents a significant investment in one of the translational steps toward improving the quality and value of health care for all (197,198). Health services research, of which CER is only a part, has been estimated to account for 1.5% of total biomedical research expenditures and 0.1% of the total U.S. expenditures

on health care (116), but the ARRA funding may reflect a trend toward increased investment in these translational building blocks for improving health. The council believes that these investments create the potential for training a new cadre of researchers, invigorating current researchers, and improving health outcomes.

Third, CER has the potential to drive high-value innovation and to enable the practice of more personalized medicine based on subgroups of patients. The goal of randomized efficacy trials is often to prove that a treatment is superior to placebo. But a more important question may be whether the intervention is better than other available interventions for specific populations and whether we can identify the subgroups of patients who will benefit the most from or are most likely to be harmed by specific interventions. The council emphasizes that CER must focus on informing the care of people who are often excluded from trials (e.g., those with multiple chronic conditions) and identifying subgroups of patients (e.g., elderly, racial, and ethnic minorities or people with a particular genetic marker) whose response to a given therapy or intervention may be different from that of the “average” patient in a trial.

5.3 Impact of CER

Table 3. *Threshold and prioritization criteria outlined by the Federal Coordinating Council for Comparative Effectiveness Research.**

<p>Minimum threshold criteria for projects (must be met for a project to be considered)</p> <ul style="list-style-type: none"> • Inclusion within statutory limits of ARRA and the Council’s definition of CER • Potential to inform decision making by patients, clinicians, or other stakeholders • Responsiveness to expressed needs of patients, clinicians, or other stakeholders • Feasibility of research topic <p>Prioritization criteria for scientifically meritorious research and investments</p> <ul style="list-style-type: none"> • Potential impact (e.g., prevalence of condition, burden of disease, variability among outcomes, costs) • Potential for evaluating comparative effectiveness among diverse populations and engaging communities in research • Addressing of uncertainty within the clinical and public health communities regarding management decisions and variability in practice • Addressing of a need or gap unlikely to be addressed through other organizations • Potential for multiplicative effect (e.g., laying of a foundation for future CER, such as data infrastructure and methods development and training, or generating of additional investment outside government)

* ARRA denotes the American Recovery and Reinvestment Act, and CER comparative-effectiveness research.

Source: Conway PH, Clancy C. Comparative-effectiveness research – implications of the Federal Coordinating Council’s Report. *N Engl J Med* 2009; 361:328-330 (196).

The projected ten year impact on spending as a result of a center for comparative effectiveness has been shown to provide approximately \$368 billion system-wide savings, with approximately \$114 billion in savings for the federal government, \$49 billion for state and local governments, approximately \$98 billion for private payors, and \$107 billion for households (Fig. 10) (199).

5.4 What is Ideal CER?

An ideal organizational structure should be at "arms length" from government, preferably through a federally chartered corporation, that remains accountable to Congress, but has stable funding and a governance structure insulated from overt political pressure (190). Thus, the organizational structure will establish a clear process for stakeholder engagement across all functions of the organization while setting the governance structure and internal politics to minimize the perceptions of bias. Second, as Pearson (190) describes, the organization structure will develop a common methodology and format for comparative effectiveness reviews, but commissions their production from a national network of academic and private sector review organizations which can produce high-quality evidence reviews and develop review rapidly, keep them up to date, and ensure that they reflect the contributions of all stakeholders.

While this structure appears to be ideal, this may also be associated with multiple flaws specifically with the bias, definition of stakeholders, methodologists, and clinicians. However, coordination by a single high-profile, trusted national CER organization, free of bias and influenced by administration, can achieve the greatest influence.

Currently CER is diffused across many sectors of the health care system. While diversity and competition have advantages, the broader impact of the lack of coordination has impaired the ability of comparative effectiveness evidence to help the health care system achieve desired goals. Consequently, a federal comparative effectiveness organization should therefore exercise a leading role in supporting methods development, establishing consensus for key methods and procedures within comparative effectiveness reviews, and developing consistent frameworks and formats for communicating results. The ultimate goal should be for the organization to be accountable to Congress even though it is structured and funded so that it has greater political insulation than existing structures inside the government.

6.0 METHODOLOGY OF CER

Tunis (200) described that effective CER will require new research methodology for reaching conclusions about the benefits, risks, and costs of actual medical

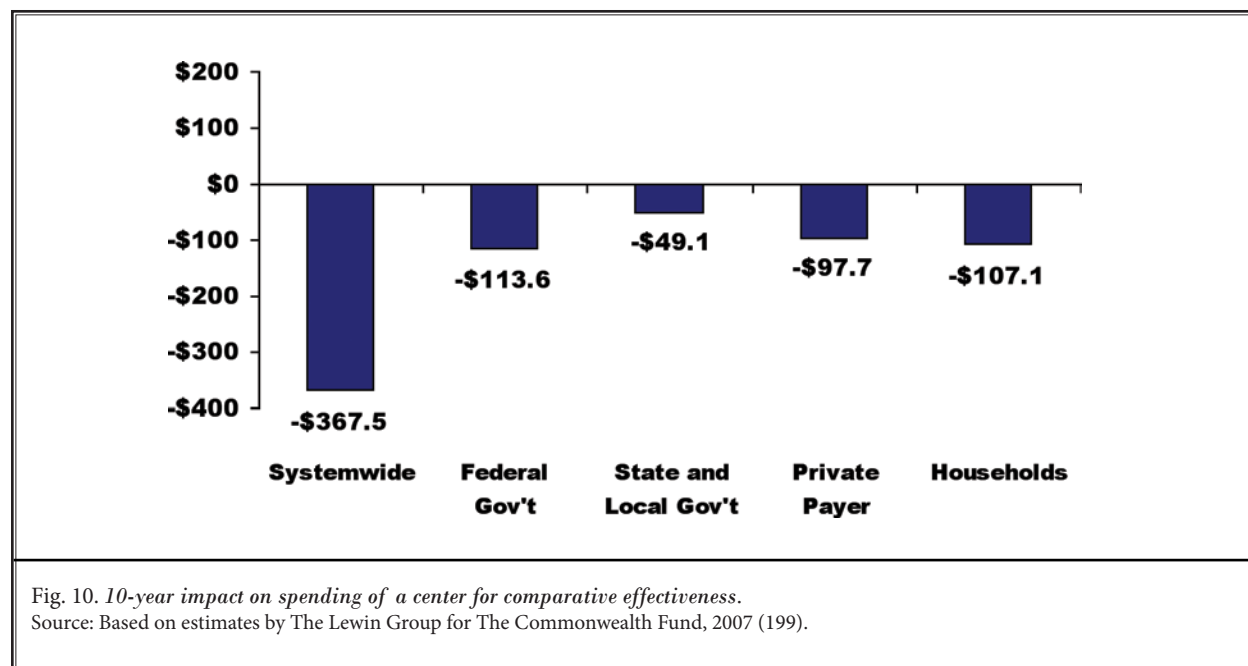


Table 4. *Hierarchy of evidence.*

CER	EBM
Systematic reviews of existing research, including meta-analysis	Systematic reviews of randomized trials
Experimental studies, including randomized clinical trials (RCTs), in which patients or groups of patients are assigned to alternative treatments, practices, or policies	Single randomized trial
Prospective non-experimental studies, including registries, which observe patterns of care and outcomes, but do not assign patients to specific study groups	Systematic review of observational studies addressing patient-important outcomes
Retrospective analysis of existing clinical or administrative data, including natural experiments	Single observational study addressing patient-important outcomes
Decision modeling, with or without cost information	Physiologic studies (studies of blood pressure, cardiac output, exercise capacity, bone density, and so forth)

Adapted and modified from Tunis SR. Strategies to improve comparative effectiveness research methods and data infrastructure. In: *The Hamilton Project. Implementing Comparative Effectiveness Research: Priorities, Methods, and Impact.* Engelberg Center for Health Care Reform at Brookings. June 2009, pp 35-54 (200).

practices, and much better data infrastructure to provide the foundation for this evidence. Thus, to achieve the multiple technical objectives, methodological guidance is the practice for the design of CER studies that reflect decision maker needs and balance internal validity with feasibility and timeliness. Further, these studies must involve patients, consumers, clinicians, payors, policy makers, and other relevant decision makers in key cases of CER study development and implementation. Finally, Tunis (200) describes the research infrastructure has to be improved to enhance the efficiency of CER studies.

The single most important factors for CER and its utilization is that CER must be valid, non-biased, relevant, timely, feasible, and actionable. To achieve these goals, the current approaches of conducting clinical and health services research may not suffice. It is essential to go beyond the current approaches of conducting clinical and health services research and to go beyond the traditional research community. Thus, the CER can be performed using a broad range of established and emerging methods. Table 4 illustrates categories of methods of CER (200) and hierarchy of evidence of EBM.

Thus, experimental studies will continue to be a crucial source of CER information, and for those questions that are best addressed with these methods, it is

critically important to develop study designs and infrastructure that will generate credible and relevant information, as quickly and inexpensively as possible (200). However, in contrast to EBM, non-experimental approaches are a useful tool for CER and continue to become increasingly important as such methods continue to be refined (201-203).

6.1 Principles of Methodology

In general, while EBM and CER go hand-in-hand, and CER can be derived from EBM as shown in Figure 9 (194), it has been stated that the “traditional hierarchies of evidence” are overly simplistic and should not necessarily guide the implementation of CER. Rawlins (204), chairman of the NICE in the United Kingdom described that “hierarchies of evidence should be replaced by accepting – indeed embracing – a diversity of approaches.” However, it has been stated that this is not a plea to abandon RCTs and replace them with observational studies, but rather it is a plea to investigators to continue to develop and improve their methods; to decision makers to avoid adapting entrenched positions about the nature of the evidence; and for both to accept that interpretation of evidence requires judgement. Thus, it re-emphasizes knowing that an intervention works under ideal circumstances (efficacy is necessary but not sufficient for evaluating what is appropriate for pa-

tients in real-world practice settings). Consequently, to obtain information on real-world practice settings, and to answer CER questions, an alternate approach to the generation and appraisal of evidence may be required (205).

Categories described in Table 4 include all types of evidence including systematic reviews of existing research, including meta-analysis as the number one category; however, in this hierarchy experimental studies, including RCTs, in which patients or groups of patients are assigned to alternative treatments, practices or policies are also important. Further, an adjunct to RCTs in the context of CER will be data collected during the delivery of and payment for health care. There also have been important advances in methods that improve the validity of analysis of non-experimental data (201). In addition, the advances made in methods to improve the validity of analysis of observational data, along with design and use of clinical registries (206), and technical advances in documentation, will improve the quality of non-experimental data.

However, the major worry is the bias, the definition of the methodology, and lack of application of appropriate methodology. As an example, the disadvantages of equivalence trials and practical trials is that many researchers, even the ones acclaimed to be the international experts and involved in decision-making at AHRQ and other organizations, failed to recognize the difference between a placebo and alternate treatment.

For example, Levin (207) and Smuck (208) consider steroid as ineffective as they showed no significant difference with local anesthetic, even though both groups experienced significant relief from baseline with functional status improvement (209-217). The basis is that they consider local anesthetic as the placebo. Many other investigations, including Cochrane reviews, have ignored this relationship and have provided misinformation. Many conclusions also are confusing for policymakers and academicians when it is stated that there was no significant difference between local anesthetic or steroid, leading to the impression that steroids do not work and local anesthetic is a placebo. Researchers of CER should put their personal biases aside and implement CER without confusing placebo control with CER.

6.3 Selection of Topics of CER

In choosing topics for CERs, a number of criteria are considered, including burden of illness; evidence suggesting underuse or overuse; the cost of the intervention or of not treating the illness; controversy sur-

rounding the treatment; and interventions intended to treat conditions that disproportionately affect women, traditionally underserved minorities, the elderly, and children. Prior to the establishment of the CER in the United States, the effective healthcare (EHC) program research, originating from the Medicare Prescription Drug, Improvement, and Modernization Act originating from the MMA of 2003 and the AHRQ, conducted research on multiple topics (156,218). The first 14 CERs were conducted from 2005 through 2007, of which 13 were therapy and one was of diagnosis (218). The comparative effectiveness study conducted for the diagnostic purposes was effectiveness of non-invasive diagnostic tests for breast abnormalities (218).

6.3 Types of Methodology

Given the multiple unsubstantiated statements about the effectiveness of CER, several questions do occur. If it is easy to obtain such data, opponents wonder why we do not have such data or we are not in the process of getting such data (101). Further, what kind of data would resolve regional practices, provide cost effectiveness, and also provide cost savings while providing efficient quality health care to all Americans. Consequently, the type of research that would best yield comparative effectiveness data is uncertain. Prospective RCTs, by eliminating or minimizing confounding variables, would provide the highest quality data (10). However, such trials are very expensive and usually lengthy. Further, these trials are small and also include a small proportion of affected population, often excluding patients with comorbidities that are prevalent in clinical practice. Registry data can overcome these limitations, but invariably introduce uncontrolled variables. Even systematic reviews may be criticized for their limitations. Thus, it seems clear that the process of acquiring accurate high-quality CER data will neither be easy nor inexpensive.

The CBO (20) describes that analyzing existing data would require a different set of skills and would cost less than overseeing new clinical trials that compared different treatments. Further, the scope of analysis – both the types of comparisons and the questions that analysis would address would differ. Consequently, federal efforts to assess different treatment options could be pursued in a variety of ways. Options range from synthesizing existing research or a systematic review, to conducting new studies using data that are already available to fund new head-to-head clinical trials. However, all options could be performed at the same time.

Systematic reviews of existing research would probably be easier to implement, which will only require a review and summary of the results of existing studies in a systematic and rigorous way. Even though existing studies may only compare a single treatment to a placebo, the results of several studies of individual therapies could in some cases be combined to measure those treatments against one another. Such reviews would be comparable to some of the work that AHRQ is already undertaking and to some current efforts based at universities or other public and private research centers, such as Economic Cycle Research Institute (ECRI). While this is promoted as a relatively low expense, in a government setting a single systematic review might cost a few hundred thousand dollars. Numerous limitations of this approach include incomplete data, lack of new information from other systematic reviews, and bias with conflicting opinions. Finally, no matter how rigorously a systematic review is conducted, its contribution is by definition constrained by the extent and quality of the underlying evidence and conflicts of interest. As an example, a systematic review of drug treatments for one form of diabetes that was sponsored by AHRQ illustrates both the strengths and weaknesses of such research (193). The review covered a large body of literature, consisting of over 200 reports, and it was able to reach a relatively clear conclusion, "Older drugs were found to be at least as effective as new drugs in controlling patient's blood sugar and cholesterol levels." The limitations included that most studies were of short duration of follow-up and the studies focused on non-elderly white patients and diabetics without comorbidities, leading to the review's recommendation that several clinical trials be conducted to fill in those gaps.

It is interesting to note that NICE relies solely on systematic reviews of available studies. NICE analyzed many different treatments on the basis of their cost-effectiveness and developed an extensive set of clinical guidelines and recommendations about using medical technologies (219) associated with substantial criticism from the public and providers. The EHC health care reports also may provide insight into new and upcoming CER. AHRQ, through their effective health care program, authorized by MMA, has published and revised 14 comparative effectiveness studies (218). They used mostly the data applicable to systematic reviews and their updates in assessing the need to update CERs. At present, the methodology utilized includes opinions from a minimum of 4 experts on each report topic, including that of the director of the EPC that conducted

the original report for a reassessment.

One of the options most discussed in recent days is medical registries. Registries collect additional information that is typically not contained in claims records, such as measures of health status or test results. Registries are essentially observational studies that involve the systematic collection of uniform baseline data and, at least in theory, provide comprehensive follow-up information for a representative population of patients. For instance, a registry might study those with a particular disease or condition, or those receiving various treatments such as a procedure, a drug, or a device. The IOM suggested that the highest priority research topic should be the creation of a prospective patient registry to compare the effectiveness of treatment strategies for low back pain without neurological deficit or spinal deformity. As one of the recommendations, the panel has recommended the 25 most important research topics in the entire comparative effectiveness field in the first quartile. The IOM panel suggested that large scale registries and other longitudinal studies would be a way of tracking the impact of medical interventions on ordinary patients in the complexities of real-world situations – and not just on the highly selected populations studied in RCTs. If we take low back pain as an example, there are more than 200 treatments for low back pain and tens of millions of patients receiving them every year in the United States alone. Registries could chart the outcomes of a representative selection of these patients who appeared to be a research project of unprecedented magnitude, even though it may appear simple. Large registries would require substantial technological infrastructure, new interactive data networks, and incentives to encourage health care institutions, health care providers, and patients to participate. In addition, the IOM panel acknowledged that there is a need for further research on how best to interpret the information provided by these large observational studies.

Skeptics point out that registries have a sparse track record in general and in back pain research specifically, are vulnerable to significant biases, and do not support firm conclusions about the effectiveness of common treatments. Carragee et al (74) called for early CER as part of a broader effort to improve spine research and emphasized that the only treatments and technologies should be widely disseminated or those supported by strong evidence. Koes (220), an author of numerous publications on low back pain and treatments expressed that, "If the question is to investigate the effectiveness of treatment, then a registry is obviously not the most valid method." He suggested that researchers carefully consider the purpose of registries

before they begin, particularly the research questions to be posed and answered. In addition, if a registry is not started with a clear focused question, they will not be useful. Deyo (221), a proponent of restraint for various types of spinal treatments and an opponent of overtreating back pain (38,222) was puzzled by the specification of a registry as a method of comparative treatment for low back pain. He suggested that the registries are a complementary rather than a stand-alone research tool. However, Deyo and Mirza (222) pointed out that registries may have a major role to play in post-marketing surveillance of new technologies. For valid information to be gathered from registries, they have to be complete and provide follow-up on nearly every patient. Chou (223), of the Oregon Health and Science University Evidence-Based Practice Center, and author of numerous evidence-based manuscripts, was also surprised at the IOM's recommendations of a registry as the highest priority back pain research project.

Registries focused on specific treatments could also be subject to bias if those patients differed systematically from patients who did not receive those treatments. However, this problem could be addressed by including a comparison group in the registries. Another concern is the data elements to collect would include a more extensive list permitting a richer analysis, but also would raise the burden of participation and again, the accuracy. Further, extensive registries and registries involving all patients to participate will be very expensive to operate, with the annual costs of maintaining a typical registry probably in the order of several million dollars (20).

The establishment of registries could affect medical practice in various ways. As an example, CMS instituted a policy of coverage with "evidence development" for Medicare, to address treatments with potentially promising but uncertain medical benefits (224). Under the policy, Medicare covers the costs of implantable cardioverter-defibrillators for a broad set of heart conditions that had previously been eligible – but only if those new patients are included in a registry that is supposed to track their progress. If CMS would otherwise have decided not to cover that treatment for those patients, then the new policy means an increase in spending in the near term, but it also allows broader access to the technology in order to help generate the kind of evidence needed to reach a conclusion about its value (20). The registry may also help ensure, through its documentation requirements, that all patients meet the medical criteria required for Medicare coverage. In fact,

in Sweden (225), health costs were reduced by avoiding repeat operations to fix faulty or poorly installed hips, with a registry of patients undergoing hip replacement surgery.

Finally, the method of research that would probably yield the most-definitive results involves RCTs to compare treatments head to head, but that approach would also be the most expensive and would take the longest to conduct. The CBO states that the total cost for conducting an extensive trial can exceed \$100 million over the course of the study, even though many trials are less expensive, and some may cost only a few million dollars (20). Further, to address many issues related to placebo-controlled trials, equivalence or non-inferiority trials have been advocated. These may be considered as practical clinical trials (PCTs) (25) with 2 key features that they compare treatment choices that clinicians face and include a wide variety of study participants drawn from a range of practice settings.

Another approach that has been suggested as an alternative or supplement to clinical trials is the use of computer models to simulate the effects of treatments on different populations of patients. While many well-designed models exist, perhaps the most prominent one is known as the Archimedes development, led by David Eddy with the support of Kaiser Permanent Health Plan (226). However, many obstacles lie in the path of modeling. Even the models rich enough to simulate real-world medical care may not be transparent enough to generate confidence in or acceptance of their results.

6.4 Updating the Methodology of CER

Good research practices for CER were recently evaluated and published (227-229). In 2007, the International Society for Pharmacoeconomics and Outcomes Research (ISPOR) Health Science Policy Council recommended the establishment of a task force to recommend good research practices for designing and analyzing retrospective databases. Two other task forces were also formed to evaluate approaches to mitigate bias and confounding in the design of non-randomized studies of treatment effects using secondary data sources and to develop analytic methods to improve causal interference from non-randomized studies of treatment effects using secondary data sources (227-229).

Part 1 of the report, defining, reporting, and interpreting non-randomized studies of treatment effects using secondary data sources (227), proposed 4 primary characteristics – relevance, specificity, novelty, and feasibility, while defining the research question. Recommen-

dations included: the practice of a priori specification of the research question; transparency of prespecified analytical plans; provision of justification for any subsequent changes in the analytical plan; reporting the results of prespecified plans as well as results from significant modifications; structured abstracts to report findings with scientific neutrality; and reasoned interpretations of findings to help inform policy decisions. The task force report also concluded that CER, in the form of non-randomized studies using secondary databases can be designed with rigorous elements and conducted with sophisticated statistical methods to improve causal inference of treatment effects.

Part II of the task force report (228), evaluating the approaches to mitigate bias and confounding in the design of non-randomized studies of treatment effects using secondary data sources for CER, provided recommendations and tools for researchers to mitigate threats to validity from bias and confounding in measurement of exposure and outcomes. Recommendations on design included: the need for a data analysis plan with causal diagrams; detailed attention to classification bias and definition of exposure and clinical outcomes; careful and appropriate use of restriction; and extreme care to identify and control for confounding factors, including time-dependent confounding. In this part they concluded that the design of non-randomized studies of comparative effectiveness face several daunting issues, including measurement of exposure and outcome challenged by misclassification and confounding.

Part III of good research practice for CER (229) described analytic methods to improve causal inferences from non-randomized studies of treatment effects using secondary data sources. The task force recommended that general analytic techniques and specific best practices where consensus is reaching include: use of stratification analysis before multivariable modeling,

multivariable regression including model performance and diagnostic testing, propensity scoring, instrumental variable, and structural modeling techniques including marginal structural models, where appropriate for secondary data. They concluded that valid findings of causal therapeutic benefits can be produced from non-randomized studies using an area of state-of-the-art analytic techniques. Further, they added that improving the quality and uniformity of these studies will improve the value to patients, physicians, and policy makers worldwide.

CONCLUSION

In Part 1 of this health policy review we have described multiple considerations of CER and the basis for its introduction. Further, we also discussed the evolution of CER in the United States as well as internationally; similarities and differences between CER and EBM; the origin, structure and role of the federal coordinating council; and methodology including principles, selection of topics of CER, types of methodology utilized, and emerging suggestions on updating the methodology of CER.

Part 2 of this series will describe the potential outcomes of CER, the impact of CER on the practice of medicine, comparison of CER in the United States and other countries where CER is well developed, an ideal CER, and finally the impact on interventional pain management.

ACKNOWLEDGMENTS

We would like to thank the editorial board of Pain Physician for review and criticism in improving the manuscript. The authors wish to thank Tonie M. Hatton and Diane E. Neihoff, transcriptionists, for their assistance in preparation of this manuscript.

REFERENCES

- Evans H. Comparative effectiveness in health care reform: Lessons from abroad. *Backgrounder* No. 2239; February 4, 2009.
- Brown SR. *Scurvy: How a Surgeon, a Mariner, and a Gentleman Solved the Greatest Medical Mystery of the Age of Sail*. St. Martin's Press, New York, 2003.
- Codman EA. *A Study in Hospital Efficiency* (Boston, Mass.: Privately printed, 1916).
- Committee on Comparative Effectiveness Research Prioritization, Institute of Medicine. *Initial National Priorities for Comparative Effectiveness Research*. National Academy of Sciences, Washington DC, 2009.
- Sackett DL, Rosenberg WM, Gray JA, Haynes RB, Richardson WS. Evidence based medicine: What it is and what it isn't. *BMJ* 1996; 312:71-72.
- Manchikanti L. Evidence-based medicine, systematic reviews, and guidelines in interventional pain management: Part 1: Introduction and general considerations. *Pain Physician* 2008; 11:161-186.
- Manchikanti L, Boswell MV, Giordano J. Evidence-based interventional pain management: Principles, problems, potential and applications. *Pain Physician* 2007; 10:329-356.
- Selker HP. Comparative effectiveness research: Medical practice, payments, and politics: The need to retain standards of medical research. *J Gen Intern Med* 2009; 24:776-778.
- Eden J, Wheatley B, McNeil B, Sox H. *Knowing What Works in Health Care: A Roadmap for the Nation*. National Academies Press, Washington, DC, 2008.
- Manchikanti L, Hirsch JA, Smith HS. Evidence-based medicine, systematic reviews, and guidelines in interventional pain management: Part 2: Randomized controlled trials. *Pain Physician* 2008; 11:717-773.
- Manchikanti L, Benyamin RM, Helm S, Hirsch JA. Evidence-based medicine, systematic reviews, and guidelines in interventional pain management: Part 3: Systematic reviews and meta-analysis of randomized trials. *Pain Physician* 2009; 12:35-72.
- Manchikanti L, Singh V, Smith HS, Hirsch JA. Evidence-based medicine, systematic reviews, and guidelines in interventional pain management: Part 4: Observational studies. *Pain Physician* 2009; 12:73-108.
- Manchikanti L, Derby R, Wolfer LR, Singh V, Datta S, Hirsch JA. Evidence-based medicine, systematic reviews, and guidelines in interventional pain management: Part 5. Diagnostic accuracy studies. *Pain Physician* 2009; 12:517-540.
- Manchikanti L, Datta S, Smith HS, Hirsch JA. Evidence-based medicine, systematic reviews, and guidelines in interventional pain management: Part 6. Systematic reviews and meta-analyses of observational studies. *Pain Physician* 2009; 12:819-850.
- Manchikanti L, Derby R, Wolfer LR, Singh V, Datta S, Hirsch JA. Evidence-based medicine, systematic reviews, and guidelines in interventional pain management: Part 7: Systematic reviews and meta-analyses of diagnostic accuracy studies. *Pain Physician* 2009; 12:929-963.
- Garber AM, Tunis SR. Does comparative effectiveness research threaten personalized medicine? *New Engl J Med* 2009; 360:1925-1927.
- Lauer MS. Comparative effectiveness research: The view from the NHLBI. *J Am Coll Cardiol* 2009; 53:1084-1086.
- Cook DA. Trojan horse reportedly seen rolling down Pennsylvania Avenue. *J Med Assoc Ga* 2009; 98:4.
- Manchikanti L, Hirsch JA. Obama health care for all Americans: Practical implications. *Pain Physician* 2009; 12:289-304.
- The Congress of the United States - Congressional Budget Office. *A CBO Paper - Research on the Comparative Effectiveness of Medical Treatments: Issues and Options for an Expanded Federal Role*. December 2007.
- National Institute for Health and Clinical Excellence (NICE). <http://www.nice.org.uk/>
- Lessons from Abroad for Health Reform in the U.S. International Policy Network and the Galen Institute, March 2009. www.galen.org/fileuploads/Lessons-FromAbroadMarch2009.pdf
- US Department of Health and Human Services. Centers for Medicare and Medicaid Services (CMS) Medicare Carriers Manual. Part 3 - Claims Process. Change Request 2337, Section 2207, Coding Physician Specialty. November 1, 2002.
- Medicare Payment Advisory Commission. *Report to the Congress: Paying for Interventional Pain Services in Ambulatory Settings*. Washington, DC: MedPAC, December 2001.
- Tunis SR, Stryer DB, Clancy CM. Practical clinical trials: Increasing the value of clinical research for decision-making in clinical and health policy. *JAMA* 2003; 290:1624-1632.
- Wennberg J, Gittelsohn A. Small area variation in health care delivery. *Sci Am* 1973; 182:1102-1108.
- Schuster MA, McGlynn EA, Brook RH. How good is the quality of health care in the United States? *Milbank Q* 1998; 76:517-563.
- Manchikanti L, Giordano J. Physician payment 2008 for interventionalists: Current state of health care policy. *Pain Physician* 2007; 10:607-626.
- Manchikanti L, Boswell MV. Interventional techniques in ambulatory surgical centers: A look at the new payment system. *Pain Physician* 2007; 10:627-650.
- Manchikanti L. Health care reform in the United States: Radical surgery needed now more than ever. *Pain Physician* 2008; 11:13-42.
- Manchikanti L, Singh V, Pampati V, Smith HS, Hirsch JA. Analysis of growth of interventional techniques in managing chronic pain in Medicare population: A 10-year evaluation from 1997 to 2006. *Pain Physician* 2009; 12:9-34.
- Buntin MB, Zuckerman S, Berenson R, Patel A, Nickols T; RAND Health - The Urban Institute. Working Paper: Volume Growth in Medicare. An Investigation of Ten Physicians' Services. Prepared for the Assistant Secretary for Planning and Evaluation, US Department of Health and Human Services. December 2008.
- Friedly J, Chan L, Deyo R. Increases in lumbosacral injections in the Medicare population: 1994 to 2001. *Spine (Phila Pa 1976)* 2007; 32:1754-1760.
- Block S, Appleby J. Employers fear cost of health insurance for jobless workers. *USA Today*, 2/20/2009.

35. Martin BI, Turner JA, Mirza SK, Lee MJ, Comstock BA, Deyo RA. Trends in health care expenditures, utilization, and health status among US adults with spine problems, 1997-2006. *Spine (Phila Pa 1976)* 2009; 34:2077-2084.
36. Deyo RA, Mirza SK, Martin BI. Back pain prevalence and visit rates: estimates from US national surveys, 2002. *Spine (Phila Pa 1976)* 2006; 31:2724-2727.
37. Martin BI, Deyo RA, Mirza SK, Turner JA, Comstock BA, Hollingworth W, Sullivan SD. Expenditures and health status among adults with back and neck problems. *JAMA* 2008; 299:656-664.
38. Deyo RA, Mirza SK, Turner JA, Martin BI. Overtreating chronic back pain: Time to back off? *J Am Board Fam Med* 2009; 22:62-68.
39. Mitchell JM. Utilization trends for advanced imaging procedures: evidence from individuals with private insurance coverage in California. *Medical Care* 2008; 46:460-466
40. Manchikanti L, Singh A. Therapeutic opioids: A ten-year perspective on the complexities and complications of the escalating use, abuse, and nonmedical use of opioids. *Pain Physician* 2008; 11: S63-S88.
41. Compton WM, Volkow ND. Major increases in opioid analgesic abuse in the United States: concerns and strategies. *Drug Alcohol Depend* 2006; 81:103-107.
42. Caudill-Slosberg MA, Schwartz LM, Woloshin S. Office visits and analgesic prescriptions for musculoskeletal pain in US: 1980 vs. 2000. *Pain* 2004; 109:514-519.
43. Franklin GM, Mai J, Wickizer T, Turner JA, Fulton-Kehoe D, Grant L. Opioid dosing trends and mortality in Washington State workers' compensation, 1996-2002. *Am J Ind Med* 2005; 48:91-99.
44. Luo X, Pietrobon R, Hey L. Patterns and trends in opioid use among individuals with back pain in the United States. *Spine (Phila Pa 1976)* 2004; 29:884-891.
45. Zerzan JT, Morden NE, Soumerai S, Ross-Degnan D, Roughead E, Zhang F, Simoni-Wastila L, Sullivan SD. Trends and geographic variation of opiate medication use in state Medicaid fee-for-service programs, 1996 to 2002. *Med Care* 2006; 44:1005-1010.
46. Deyo RA, Gray DT, Kreuter W, Mirza S, Martin BI. United States trends in lumbar fusion surgery for degenerative conditions. *Spine (Phila Pa 1976)* 2005; 30:1441-1445.
47. Martin BI, Mirza SK, Comstock BA, Gray DT, Kreuter W, Deyo RA. Are lumbar spine reoperation rates falling with greater use of fusion surgery and new surgical technology? *Spine (Phila Pa 1976)* 2007; 32:2119-2126.
48. Lad SP, Patil CG, Berta S, Santarelli JG, Ho C, Boakye M. National trends in spinal fusion for cervical spondylotic myelopathy. *Surg Neurol* 2009; 71:66-69.
49. Wennberg JE, Fisher ES, Skinner JS. Geography and the debate over Medicare reform. *Health Aff (Millwood)* 2002; Suppl Web Exclusives:W96-114.
50. Wennberg JE, Fisher ES, Skinner JS. *The Dartmouth Atlas of Health Care*. Heath Forum Inc, Dartmouth, 1999.
51. Wennberg JE, Birkmeyer JD, Birkmeyer NJO. *The Dartmouth Atlas of Cardiovascular Health Care*. AHA Press, Chicago, 1999.
52. Stukel TA, Lucas FL, Wennberg DE. Long-term outcomes of regional variations in intensity of invasive vs medical management of Medicare patients with acute myocardial infarction. *JAMA* 2005; 293:1329-1337.
53. Skinner JS, Staiger DO, Fisher ES. Is technological change in medicine always worth it? The case of acute myocardial infarction. *Health Aff (Millwood)* 2006; 25:w34-w47.
54. www.dartmouthatlas.org.
55. Cutler D, Sheiner L. The geography of Medicare. *Am Econ Rev* 1999; 89:228-233.
56. Swedlow A, Johnson G, Smithline N, Milstein A. Increased costs and rates of use in the California workers' compensation system as a result of self-referral by physicians. *N Engl J Med* 1992; 327:1502-1506.
57. Schroth WS, Schectman JM, Elinsky EG, Panagides JC. Utilization of medical services for the treatment of acute low back pain: Conformance with clinical guidelines. *J Gen Intern Med* 1992; 7:486-491.
58. Keller RB, Atlas SJ, Soule DN, Singer DE, Deyo RA. Relationship between rates and outcomes of operative treatment for lumbar disc herniation and spinal stenosis. *J Bone Joint Surg* 1999; 81:752-762.
59. Friedly J, Chan L, Deyo R. Geographic variation in epidural steroid injection use in Medicare patients. *J Bone Joint Surg Am* 2008; 90:1730-1737.
60. Doyle Jr. JJ. Returns to Local-Area Health Care Spending: Using Health Shocks to Patients Far From Home. NBER Working Paper 13301 (National Bureau of Economic Research, August 2007).
61. Fisher ES, Wennberg DE, Stukel TA, Gottlieb DJ, Lucas FL, Pinder EL. The implications of regional variations in Medicare spending, Part 2: Health Outcomes and Satisfaction with Care. *Ann Intern Med* 2003; 138:288-298.
62. Kuklina EV, Yoon PW, Keenan NL. Trends in high levels of low-density lipoprotein cholesterol in the United States, 1999-2006. *JAMA* 2009; 302:2104-2110.
63. Vogt MT, Kwok CK, Cope DK, Osial TA, Culyba M, Starz TW. Analgesic usage for low back pain: Impact on health care costs and service use. *Spine (Phila Pa 1976)* 2005; 30:1075-1081.
64. Emanuel EJ. What cannot be said on television about health care. *JAMA* 2007; 297:2131-2133.
65. Glover JA. The Incidence of Tonsillectomy in School Children. *Proc R Soc Med* 1938; 31:1219-1236.
66. Sanders D, Coulter A, McPherson K. *Variations in Hospital Admission Rates: A Review of the Literature*. King Edward's Hospital Fund for London, London, 1989, p 31.
67. Cochrane A. *Effectiveness and Efficiency: Random Reflections on Health Services* Nuffield Provincial Hospitals Trust, Leeds, 1972.
68. Wennberg JE, Mulley AG Jr, Hanley D, Timothy RP, Fowler FJ Jr, Roos NP, Barry MJ, McPherson K, Greenberg ER, Soule D, Bubolz T, Fisher E, Malenka D. An assessment of prostatectomy for benign urinary tract obstruction: Geographic variations and the evaluation of medical care outcomes. *JAMA* 1988; 259:3027-3030.
69. Partridge AH, Winer EP. On mammography - more agreement than disagreement. *N Eng J Med* 2009; 361:2499-2501.
70. The Cochrane Collaboration. *Evidence Health Care*. www.cochrane.org/docs/ebm.htm.
71. The Hamilton Project. *Implementing Comparative Effectiveness Research: Priorities, Methods, and Impact*. Engelberg Center for Health Care Reform at

- Brookings. June 2009.
72. Sung NS, Crowley WF Jr, Genel M, Salber P, Sandy L, Sherwood LM, Johnson SB, Catanese V, Tilson H, Getz K, Larson EL, Scheinberg D, Reece EA, Slavkin H, Dobs A, Grebb J, Martinez RA, Korn A, Rimoin D. Central challenges facing the national clinical research enterprise. *JAMA* 2003; 289:1278-1287.
 73. Institute of Medicine. *Learning What Works Best: The Nation's Need for Evidence on Comparative Effectiveness in Health Care*. Institute of Medicine of the National Academies, September 2007.
 74. Carragee EJ, Deyo RA, Kovacs FM, Peul WC, Lurie JD, Urrútia G, Corbin TP, Schoene ML. Clinical research: Is the spine field a mine field? *Spine (Phila Pa 1976)* 2009; 34:423-430.
 75. Esserman L, Shieh Y, Thompson I. Re-thinking screening for breast cancer and prostate cancer. *JAMA* 2009; 302:1685-1692.
 76. Gaziano JM, Gaziano TA. Simplifying the approach to the management of dyslipidemia. *JAMA* 2009; 302:2148-2149
 77. Sox HC, Greenfield S. Comparative effectiveness research: A report from the Institute of Medicine. *Ann Intern Med* 2009; 151:203-205.
 78. Don AS, Carragee E. A brief overview of evidence-informed management of chronic low back pain with surgery. *Spine J* 2008; 8:258-265.
 79. Chou R, Baisden J, Carragee EJ, Resnick DK, Shaffer WO, Loeser JD. Surgery for low back pain: A review of the evidence for an American Pain Society Clinical Practice Guideline. *Spine (Phila Pa 1976)* 2009; 34:1094-1109.
 80. Chou R, Loeser JD, Owens DK, Rosenquist RW, Atlas SJ, Baisden J, Carragee EJ, Grabojs M, Murphy DR, Resnick DK, Stanos SP, Shaffer WO, Wall EM; American Pain Society Low Back Pain Guideline Panel. Interventional therapies, surgery, and interdisciplinary rehabilitation for low back pain: An evidence-based clinical practice guideline from the American Pain Society. *Spine (Phila Pa 1976)* 2009; 34:1066-1077.
 81. Bigos SJ, Boyer OR, Braen GR, Brown K, Deyo R. *Acute Low Back Problems in Adults*. Clinical Practice Guideline Number 4. AHCPR Publication No. 95-0642. Agency for Health Care Policy and Research, Public Health Service, US Department of Health and Human Services, Rockville, December 1994.
 82. Boswell MV, Trescot AM, Datta S, Schultz DM, Hansen HC, Abdi S, Sehgal N, Shah RV, Singh V, Benyamin RM, Patel VB, Buenaventura RM, Colson JD, Corder HJ, Epter RS, Jasper JF, Dunbar EE, Atluri SL, Bowman RC, Deer TR, Swicegood JR, Staats PS, Smith HS, Burton AW, Kloth DS, Giordano J, Manchikanti L. Interventional techniques: Evidence-based practice guidelines in the management of chronic spinal pain. *Pain Physician* 2007; 10:7-111.
 83. Manchikanti L, Boswell MV, Singh V, Benyamin RM, Fellows B, Abdi S, Buenaventura RM, Conn A, Datta S, Derby R, Falco FJE, Erhart S, Diwan S, Hayek SM, Helm S, Parr AT, Schultz DM, Smith HS, Wolfer LR, Hirsch JA. Comprehensive evidence-based guidelines for interventional techniques in the management of chronic spinal pain. *Pain Physician* 2009; 12:699-802.
 84. *Evidence-based Practice Centers Partner's Guide*. Prepared for Agency for Healthcare Research and Quality by the EPC Coordinating Center. January 2005. www.ahrq.gov/Clinic/epcpartner/epcpartner.pdf
 85. Organisation for Economic and Cooperation and Development (OECD). *Health at a Glance 2009: OECD Indicators*, OECD Publishing, 2009. DOI: http://dx.doi.org/10.1787/health_glance-2009-en
 86. Centers for Medicare and Medicaid Services, Office of the Actuary, National Health Statistics Group, at <http://www.cms.hhs.gov/NationalHealthExpendData/> (Historical data from NHE summary including share of GDP, CY 1960-2008, file nhegdp08.zip.
 87. Hartman M, Martin A, Nuccio O, Catlin A; National Health Expenditure Accounts Team. Health spending growth at a historic low in 2008. *Health Aff (Millwood)* 2010; 29:147-155.
 88. Hartman M, Martin A, McDonnell P, Catlin A, National Health Expenditure Accounts Team. National health spending in 2007: slower drug spending contributes to lowest rate of overall growth since 1998. *Health Aff (Millwood)* 2009; 28:246-261.
 89. Congressional Budget Office. *The Long-Term Outlook for Health Care Spending*. November 2007. www.cbo.gov/ftpdocs/87xx/doc8758/11-13-LT-Health.pdf
 90. Congressional Budget Office. *The Long-Term Budget Outlook*. December 2007. www.cbo.gov/ftpdocs/88xx/doc8877/12-13-LTBO.pdf
 91. Kaiser Family Foundation, based on Mathematica Policy Research, Inc. Tracking Medicare Health and Prescription Drug Plans Monthly Report December 1999-2007. CMS Medicare Advantage, Cost, PACE, Demo, and Prescription Drug Plan Contract Report, Monthly Summary Report, July 2008.
 92. Kaiser Family Foundation/Health Research and Educational Trust. Employer Health Benefits 2009 Annual Survey. <http://ehbs.kff.org/pdf/2009/7936.pdf>
 93. Hartzband P, Groopman J. Money and the changing culture of medicine. *N Engl J Med* 2009; 360:101-103.
 94. Trescot AM, Helm S, Hansen H, Benyamin R, Glaser SE, Adlaka R, Patel S, Manchikanti L. Opioids in the management of chronic non-cancer pain: An update of American Society of the Interventional Pain Physicians' (ASIPP) guidelines. *Pain Physician* 2008; 11:S5-S62.
 95. Manchikanti L, Atluri S, Trescot AM, Giordano J. Monitoring opioid adherence in chronic pain patients: Tools, techniques and utility. *Pain Physician* 2008; 11:S155-S180.
 96. Trescot AM, Glaser SE, Hansen H, Benyamin R, Patel S, Manchikanti L. Effectiveness of opioids in the treatment of chronic non-cancer pain. *Pain Physician* 2008; 11:S181-S200.
 97. Paulozzi LJ, Budnitz DS, Xi Y. Increasing deaths from opioid analgesics in the United States. *Pharmacoepidemiol Drug Saf* 2006; 15:618-627.
 98. Martell BA, O'Connor PG, Kerns RD, Becker WC, Morales KH, Kosten TR, Fiellin DA. Systematic review: Opioid treatment for chronic back pain: Prevalence, efficacy, and association with addiction. *Ann Intern Med* 2007; 146:116-127.
 99. Lurie JD, Weinstein JN. Shared decision-making and the orthopaedic workforce. *Clin Orthop* 2001; 385:68-75.
 100. Department of Health and Human Services. Office of Inspector General (OIG). Medicare Payments for Facet Joint Injection Services (OEI-05-07-00200). September 2008. www.oig.hhs.gov/oei/reports/oei-05-07-00200.pdf
 101. DeMaria AN. Comparative effectiveness research. *J Am Coll Cardiol* 2009;

- 53:634-635.
102. Fisher ES, Wennberg JE. Health care quality, geographic variations, and the challenge of supply-sensitive care. *Perspect Biol Med* 2003; 46: 69-79.
 103. Murillo H, Reece EA, Snyderman R, Sung NS. Meeting the challenges facing clinical research: Solutions proposed by leaders of medical specialty and clinical research societies. *Acad Med* 2006; 81:107-112.
 104. Lurie JD, Birkmeyer NJ, Weinstein JN. Rates of advanced spinal imaging and spine surgery. *Spine (Phila Pa 1976)* 2003; 28:616-620.
 105. Rao JK, Kroenke K, Mihaliak KA, Eckert GJ, Weinberger M. Can guidelines impact the ordering of magnetic resonance imaging studies by primary care providers for low back pain? *Am J Manag Care* 2002; 8:27-35.
 106. Deyo RA, Mirza SK. Trends and variations in the use of spine surgery. *Clin Orthop Relat Res* 2006; 443:139-146.
 107. Irwin ZN, Hilibrand A, Gustavel M, McLain R, Shaffer W, Myers M, Glaser J, Hart RA. Variation in surgical decision making for degenerative spinal disorders. Part I: Lumbar spine. *Spine (Phila Pa 1976)* 2005; 30:2208-2213.
 108. Irwin ZN, Hilibrand A, Gustavel M, McLain R, Shaffer W, Myers M, Glaser J, Hart RA. Variation in surgical decision making for degenerative spinal disorders. Part II: Cervical spine. *Spine (Phila Pa 1976)* 2005; 30:2214-2219.
 109. Angevine PD, Arons RR, McCormick PC. National and regional rates of variation of cervical discectomy with and without anterior fusion, 1990-1999. *Spine (Phila Pa 1976)* 2003; 28:931-999.
 110. Katz JN. Lumbar spinal fusion. Surgical rates, costs, and complications. *Spine (Phila Pa 1976)* 1995; 20:78S-83S.
 111. Lieberman IH. Disc bulge bubble: Spine economics 101. *Spine J* 2004; 4:609-613.
 112. McCrory DC, Turner DA, Patwardhan MB, Richardson WL. Spinal fusion for degenerative disc disease affecting the lumbar spine (draft evidence report/technology review prepared for the Medicare Coverage Advisory Committee meeting), November, 1, 2006; www.cms.hhs.gov/determinationprocess/downloads/id41ta.pdf.
 113. Weinstein JN, Bronner KK, Morgan TS, Wennberg JE. Trends and geographic variations in major surgery for degenerative diseases of the hip, knee, and spine. *Health Aff (Millwood)* 2004; Suppl Web Exclusives:VAR81-89.
 114. Tricoci P, Allen JM, Kramer JM, Califf RM, Smith SC. Scientific evidence underlying the ACC/AHA clinical practice guidelines. *J Am Med Assoc* 2009; 301:831-841.
 115. Harpole LH, Kelley MJ, Schreiber G, Toloza EM, Kolimaga J, McCrory DC. Assessment of the scope and quality of clinical practice guidelines in lung cancer. *Chest* 2003; 123:7S-20S.
 116. Moses H 3rd, Dorsey ER, Matheson DH, Thier SO. Financial anatomy of biomedical research. *JAMA* 2005; 294:1333-1342.
 117. AcademyHealth Report. Placement, Coordination, and Funding of Health Services Research within the Federal Government. AcademyHealth, September 2005.
 118. Chou R, Atlas SJ, Stanos SP, Rosenquist RW. Nonsurgical interventional therapies for low back pain: A review of the evidence for an American Pain Society clinical practice guideline. *Spine (Phila Pa 1976)* 2009; 34:1078-1093.
 119. Chou R. Same trials, different conclusions: Sorting out discrepancies between reviews on interventional procedures of the spine. *Spine J* 2009; 9:679-689.
 120. American College of Occupational and Environmental Medicine (ACOEM) Low back Disorders. In *Occupational Medicine Practice Guidelines: Evaluation and Management of Common Health Problems and Functional Recovery of Workers*, Second Edition. American College of Occupational and Environmental Medicine Press, Elk Grove Village, 2007.
 121. American College of Occupational and Environmental Medicine (ACOEM) Chronic Pain. In *Occupational Medicine Practice Guidelines: Evaluation and Management of Common Health Problems and Functional Recovery of Workers*, Second Edition. American College of Occupational and Environmental Medicine Press, Elk Grove Village, 2008.
 122. Dennison PL, Kennedy CW. *Official Disability Guidelines*. 15th ed. Work Loss Data Institute, 2010.
 123. HAYES, Inc. Independent Health Technology Assessment Company. www.hayesinc.com
 124. Manchikanti L, Singh V, Derby R, Helm S, Trescot AM, Staats PS, Prager JP, Hirsch JA. Review of occupational medicine practice guidelines for interventional pain management and potential implications. *Pain Physician* 2008; 11:271-289.
 125. Manchikanti L, Singh V, Helm S, Trescot AM, Hirsch JA. A critical appraisal of 2007 American College of Occupational and Environmental Medicine (ACOEM) practice guidelines for interventional pain management: An independent review utilizing AGREE, AMA, IOM, and other criteria. *Pain Physician* 2008; 11:291-310.
 126. Manchikanti L, Singh V, Derby R, Schultz DM, Benyamin RM, Prager JP, Hirsch JA. Reassessment of evidence synthesis of occupational medicine practice guidelines for interventional pain management. *Pain Physician* 2008; 11:393-482.
 127. Shaneyfelt TM, Mayo-Smith MF, Rothwangl J. Are guidelines following guidelines? The methodological quality of clinical practice guidelines in the peer reviewed medical literature. *JAMA* 1999; 281:1900-1905.
 128. Staal JB, de Bie RA, de Vet HC, Hildebrandt J, Nelemans P. Injection therapy for subacute and chronic low back pain: An updated Cochrane review. *Spine (Phila Pa 1976)* 2009; 34:49-59.
 129. Airaksinen O, Brox JJ, Cedraschi C, Hildebrandt J, Kluber-Moffett J, Kovacs F, Mannion AF, Reis S, Staal JB, Ursin H, Zanoli G. Chapter 4: European guidelines for the management of chronic nonspecific low back pain. *Eur Spine J* 2006; 15:S192-S300.
 130. Boswell MV, Shah RV, Everett CR, Sehgal N, McKenzie-Brown AM, Abdi S, Bowman RC, Deer TR, Datta S, Colson JD, Spillane WF, Smith HS, Lucas-Levin LF, Burton AW, Chopra P, Staats PS, Wasserman RA, Manchikanti L. Interventional techniques in the management of chronic spinal pain: Evidence-based practice guidelines. *Pain Physician* 2005; 8:1-47.
 131. Falco FJE, Erhart S, Wargo BW, Bryce DA, Atluri S, Datta S, Hayek SM. Systematic review of diagnostic utility and therapeutic effectiveness of cervical facet joint interventions. *Pain Physician* 2009; 12:323-344.
 132. Datta S, Lee M, Falco FJE, Bryce DA, Hayek SM. Systematic assessment of diagnostic accuracy and therapeutic utility of lumbar facet joint interven-

- tions. *Pain Physician* 2009; 12:437-460.
133. Singh V, Manchikanti L, Shah RV, Dunbar EE, Glaser SE. Systematic review of thoracic discography as a diagnostic test for chronic spinal pain. *Pain Physician* 2008; 11:631-642.
 134. Wolfer L, Derby R, Lee JE, Lee SH. Systematic review of lumbar provocation discography in asymptomatic subjects with a meta-analysis of false-positive rates. *Pain Physician* 2008; 11:513-538.
 135. Manchikanti L, Dunbar EE, Wargo BW, Shah RV, Derby R, Cohen SP. Systematic review of cervical discography as a diagnostic test for chronic spinal pain. *Pain Physician* 2009; 12:305-321.
 136. Manchikanti L, Glaser S, Wolfer L, Derby R, Cohen SP. Systematic review of lumbar discography as a diagnostic test for chronic low back pain. *Pain Physician* 2009; 12:541-559.
 137. Conn A, Buenaventura R, Datta S, Abdi S, Diwan S. Systematic review of caudal epidural injections in the management of chronic low back pain. *Pain Physician* 2009; 12:109-135.
 138. Parr AT, Diwan S, Abdi S. Lumbar interlaminar epidural injections in managing chronic low back and lower extremity pain: A systematic review. *Pain Physician* 2009; 12:163-188.
 139. Benyamin RM, Singh V, Parr AT, Conn A, Diwan S, Abdi S. Systematic review of the effectiveness of cervical epidurals in the management of chronic neck pain. *Pain Physician* 2009; 12:137-157.
 140. Buenaventura RM, Datta S, Abdi S, Smith HS. Systematic review of therapeutic lumbar transforaminal epidural steroid injections. *Pain Physician* 2009; 12:233-251.
 141. Helm S, Hayek S, Benyamin RM, Manchikanti L. Systematic review of the effectiveness of thermal annular procedures in treating discogenic low back pain. *Pain Physician* 2009; 12:207-232.
 142. Smith HS, Chopra P, Patel VB, Frey ME, Rastogi R. Systematic review on the role of sedation in diagnostic spinal interventional techniques. *Pain Physician* 2009; 12:195-206.
 143. Frey ME, Manchikanti L, Benyamin RM, Schultz DM, Smith HS, Cohen SP. Spinal cord stimulation for patients with failed back surgery syndrome: A systematic review. *Pain Physician* 2009; 12:379-397.
 144. Epter RS, Helm S, Hayek SM, Benyamin RM, Smith HS, Abdi S. Systematic review of percutaneous adhesiolysis and management of chronic low back pain in post lumbar surgery syndrome. *Pain Physician* 2009; 12:361-378.
 145. Patel VB, Manchikanti L, Singh V, Schultz DM, Hayek SM, Smith HS. Systematic review of intrathecal infusion systems for long-term management of chronic non-cancer pain. *Pain Physician* 2009; 12:345-360.
 146. Rupert MP, Lee M, Manchikanti L, Datta S, Cohen SP. Evaluation of sacroiliac joint interventions: A systematic appraisal of the literature. *Pain Physician* 2009; 12:399-418.
 147. Hayek SM, Helm S, Benyamin RM, Singh V, Bryce DA, Smith HS. Effectiveness of spinal endoscopic adhesiolysis in post lumbar surgery syndrome: A systematic review. *Pain Physician* 2009; 12:419-435.
 148. Hirsch JA, Singh V, Falco FJE, Benyamin RM, Manchikanti L. Automated percutaneous lumbar discectomy for the contained herniated lumbar disc: A systematic assessment of evidence. *Pain Physician* 2009; 12:601-620.
 149. Singh V, Manchikanti L, Benyamin RM, Helm S, Hirsch JA. Percutaneous lumbar laser disc decompression: A systematic review of current evidence. *Pain Physician* 2009; 12:573-588.
 150. Singh V, Benyamin RM, Datta S, Falco FJE, Helm S, Manchikanti L. Systematic review of percutaneous lumbar mechanical disc decompression utilizing Dekompressor. *Pain Physician* 2009; 12:589-599.
 151. Manchikanti L, Derby R, Benyamin RM, Helm S, Hirsch JA. A systematic review of mechanical lumbar disc decompression with nucleoplasty. *Pain Physician* 2009; 12:561-572.
 152. Webster BS, Verma SK, Gatchel RJ. Relationship between early opioid prescribing for acute occupational low back pain and disability duration, medical costs, subsequent surgery, and late opioid use. *Spine (Phila Pa 1976)* 2007; 32:2127-2132.
 153. Lawrence JT, London N, Bohlman HH, Chin KR. Preoperative narcotic use as a predictor of clinical outcome. *Spine (Phila Pa 1976)* 2008; 33:2074-2078.
 154. Fillingim RB, Doleys DM, Edwards RR, Lowery D. Clinical characteristics of chronic back pain as a function of gender and oral opioid use. *Spine (Phila Pa 1976)* 2003; 28:143-150.
 155. Reducing the Costs of Poor-Quality Health Care through Responsible Purchasing Leadership. Chicago, IL: Midwest Business Group on Health, 2003. www.mbg.org/templates/UserFiles/Files/COPQ/copq%202nd%20printing.pdf
 156. H.R. 1. The Medicare Prescription Drug, Improvement, and Modernization Act of 2003, P.L. 108-173, Enacted December 8, 2003.
 157. Stevens A, Milne R, Burls A. Health Technology Assessment: History and Demand. *J Public Health Med* 1998; 25:98-101.
 158. Perry S. The brief life of the National Center for Healthcare Technology. *New Engl J Med* 1982; 307:1095-1100.
 159. Office of Technology Assessment. *Identifying Health Technologies that Work: Searching for Evidence*, OTA-H-608. September 1994.
 160. U.S. Department of Health and Human Services, Agency for Healthcare Research and Quality. www.ahrq.gov
 161. Gray BH, Gusmano MK, Collins SR. AHCP and the changing politics of health services research. *Health Aff (Milwood)* 2003; W3:283-307.
 162. Gonzalez EG. Preface. In Gonzalez ER, Materson RS (eds). *The Nonsurgical Management of Acute Low Back Pain*. Demos Vermande, New York, 1997, pp vii-viii.
 163. Gonzalez EG, Materson RS. The guidelines, the controversy, the book. In Gonzalez ER, Materson RS (eds). *The Nonsurgical Management of Acute Low Back Pain*. Demos Vermande, New York, 1997, pp 1-4.
 164. Jarrett A.T. Effect of the Medicare Prescription Drug Improvement and Modernization Act on the bottom line. *Am J Health Syst Pharm* 2006; 63:S10-S13.
 165. Wilensky GR. Implementing the MMA. *Healthc Financ Manage* 2004; 58:30,32.
 166. Agency for Healthcare Research and Quality. Effective Health Care Program. www.effectivehealthcare.ahrq.gov
 167. Public Law No: 111-5. H.R. 1. American Recovery and Reinvestment Act of 2009 signed by President Barack Obama on 2/17/2009.
 168. United States Department of Health and Human Services. A new era of re-

- sponsibility. February 4, 2009. www.georgiavoices.org/storage/voices/
169. Baucus MS. 3408: Comparative Effectiveness Research Act of 2008. Library of Congress, 2008.
 170. International Network of Agencies for Health Technology Assessments (INAHTA). www.inahta.org
 171. Evans H. *Sixty Years On--Who Cares for the NHS?* Institute of Economic Affairs, London, 2008, pp 26-54.
 172. Pugatch MP, Ficaí F. *A Healthy Market? An Introduction to Health Technology Assessment*. Stockholm Network, London, 2007.
 173. Velasco-Garrido M, Busse R. *Health Technology Assessment: An Introduction to Objectives, Role of Evidence, and Structure in Europe*. WHO European Observatory on Health Systems and Policies, Brussels, Belgium, 2005.
 174. Kristensen FB. EUnetHTA and health policy-making in Europe. *Eurohealth* 2006; 12:36-38.
 175. National Board of Health www.sst.dk
 176. Oliver A, Mossialos E, Robinson R. Health technology assessment and its influence on health-care priority setting. *Int J Technol Assess Health Care* 2004; 20:1-10.
 177. Draborg E, Gyrd-Hansen D. Time-trends in health technology assessments: An analysis of developments in composition of international health technology assessments from 1989 to 2002. *Int J Technol Assess Health Care* 2005; 21:492-498.
 178. García-Altés A, Ondategui-Parra S, Neumann PJ.. Cross-national comparison of technology assessment processes *Int J Technol Assess Health Care* 2004; 20:300-310.
 179. Eden J, Wheatley B, McNeil B, Sox H. Building a foundation for knowing what works in health care. In: *Knowing What Works in Health Care: A Roadmap for the Nation*. National Academies Press, Washington, DC, 2008, pp 153-178.
 180. Australian Safety & Efficacy Register of New Interventional Procedures--Surgical. 2005. *Annual Report*. Melbourne, Australia: Royal Australian College of Surgeons.
 181. CADTH (Canadian Agency for Drugs and Technologies in Health). 2006. *Health technology assessment*. www.cadth.ca/index.php/en/hta/
 182. Canadian Task Force on Preventive Health Care. 2005. *Evidence-based clinical prevention*. www.ctfphc.org
 183. Department on Health and Ageing. 2006. *About us: Our role*. www.health.gov.au/internet/wcms/publishing.nsf/Content/health-overview.htm
 184. Haute Autorité de Santé. 2007. *About HAS*. www.has-sante.fr/portail/display.jsp?id=c_5443&pcid=c_5443
 185. Institute for Quality and Efficiency in Health Care. 2007. *About us*. www.iqwig.de/about-us.21.en.html
 186. National Board of Health. 2007. *Danish Centre for Health Technology Assessment*. www.sst.dk/Planlaegning_og_behandling/Medicinsk_teknologivurdering.aspx?lang=en
 187. NICE (National Institute for Health and Clinical Excellence). 2007. *About technology appraisals*. www.nice.org.uk/page.aspx?o=202425
 188. SIGN (Scottish Intercollegiate Guidelines Network). 2007. *Guideline Development Programme*. www.sign.ac.uk/guidelines/development/index.html
 189. National Health and Medical Research Council. 2006. *Role of the NHMRC*. www.nhmrc.gov.au/about/role/index.htm
 190. Pearson S. From better evidence to better care: Using comparative effectiveness research to guide practice and policy. In: *The Hamilton Project. Implementing Comparative Effectiveness Research: Priorities, Methods, and Impact*. Engelberg Center for Health Care Reform at Brookings. June 2009, pp 55-82.
 191. Smith GC, Pell JP. Parachute use to prevent death and major trauma related to gravitational challenge: Systematic review of randomised controlled trials. *BMJ* 2003; 327:1459-1461
 192. Chassin MR. Is health care ready for Six Sigma quality? *Milbank Q* 1998; 76:565-591, 510.
 193. Bolen S, Feldman L, Vassy J, Wilson L, Yeh HC, Marinopoulos S, Wiley C, Selvin E, Wilson R, Bass EB, Brancati FL. Systematic review: Comparative effectiveness and safety of oral medications for type 2 diabetes mellitus. *Ann Intern Med* 2007; 146:386-399.
 194. Drummond MF, Schwartz JS, Jönsson B, Luce BR, Neumann PJ, Siebert U, Sullivan SD. Key principles for the improved conduct of health technology assessments for resource allocation decisions. *Int J Technol Assess Health Care* 2008; 24:244-258.
 195. Daschle T, Greenberger SS, Lambrew JM. *Critical: What We Can Do About the Health-Care Crisis*. Thomas Dunne Books, New York, 2008.
 196. Conway PH, Clancy C. Comparative-effectiveness research – implications of the Federal Coordinating Council's Report. *N Engl J Med* 2009; 361:328-330.
 197. Federal Coordinating Council for Comparative Effectiveness Research. Report to the President and Congress. Washington, DC: Department of Health and Human Services, June 2009. www.hhs.gov/recovery/programs/ce/cer-annualrpt.pdf
 198. Dougherty D, Conway PH. The “3T’s” road map to transform US health care: The “how” of high-quality care. *JAMA* 2008; 299: 2319-2321.
 199. The Lewin Group for the Commonwealth Fund, 2007.
 200. Tunis SR. Strategies to improve comparative effectiveness research methods and data infrastructure. In: *The Hamilton Project. Implementing Comparative Effectiveness Research: Priorities, Methods, and Impact*. Engelberg Center for Health Care Reform at Brookings. June 2009, pp 35-54.
 201. Schneeweiss S, Seeger JD, Landon J, Walker AM. Aprotinin during coronary-artery bypass grafting and risk of death. *NEJM* 2008; 358:771-783.
 202. Schneeweiss S, Solomon DH, Wang PS, Brookhart MA. Simultaneous assessment of short-term gastrointestinal benefits and cardiovascular risks of selective COX-2 inhibitors and non-selective NSAIDs: an instrumental variable analysis. *Arthritis Rheum* 2006; 54:3390-3398.
 203. Institute of Medicine. *The Learning Healthcare System: Workshop Summary*. Washington, DC: The National Academies Press. April 2, 2007.
 204. Rawlins M. Harveian Oration: on the evidence for decisions about the use of therapeutic interventions. *The Lancet* 2008; 372:2152-2162.
 205. Atkins D. Creating and synthesizing evidence with decision makers in mind, integrating evidence from clinical trials and other study designs. *Med Care* 2007; 45:S16-S22.
 206. Gliklich RE, Dreyer NA, eds. *Registries for Evaluating Patient Outcomes: A User's Guide*. (Prepared by Outcome DE-

- cIDE Center [Outcome Sciences, Inc. dba Outcome] under Contract No. HH-SA290200500351 TO1.) AHRQ Publication No. 07-EHC001-1. Rockville, MD: Agency for Healthcare Research and Quality. April 2007.
207. Levin JH. Prospective, double-blind, randomized placebo-controlled trials in interventional spine: What the highest quality literature tells us. *Spine J* 2009; 9:690-703.
 208. Smuck M, Levin JH. Re: Manchikanti L, Singh V, Falco FJ, et al. Cervical medial branch blocks for chronic cervical facet joint pain: A randomized, double-blind, controlled trial with one-year follow-up. *Spine* 2008; 33:1813-1820. *Spine (Phila Pa 1976)* 2009; 34:1116; author reply 1116-1117.
 209. Manchikanti L, Cash KA, McManus CD, Pampati V, Smith HS. Preliminary results of randomized, equivalence trial of fluoroscopic caudal epidural injections in managing chronic low back pain: Part 1. Discogenic pain without disc herniation or radiculitis. *Pain Physician* 2008; 11:785-800.
 210. Manchikanti L, Singh V, Cash KA, Pampati V, Damron KS, Boswell MV. Preliminary results of randomized, equivalence trial of fluoroscopic caudal epidural injections in managing chronic low back pain: Part 2. Disc herniation and radiculitis. *Pain Physician* 2008; 11:801-815.
 211. Manchikanti L, Singh V, Cash KA, Pampati V, Datta S. Preliminary results of randomized, equivalence trial of fluoroscopic caudal epidural injections in managing chronic low back pain: Part 3. Post surgery syndrome. *Pain Physician* 2008; 11:817-831.
 212. Manchikanti L, Cash KA, McManus CD, Pampati V, Abdi S. Preliminary results of randomized, equivalence trial of fluoroscopic caudal epidural injections in managing chronic low back pain: Part 4. Spinal stenosis. *Pain Physician* 2008; 11:833-848.
 213. Manchikanti L, Singh V, Falco FJE, Cash KA, Pampati V. Effectiveness of thoracic medial branch blocks in managing chronic pain: A preliminary report of a randomized, double-blind controlled trial; Clinical trial NCT00355706. *Pain Physician* 2008; 11:491-504.
 214. Manchikanti L, Singh V, Falco FJ, Cash KA, Fellows B. Cervical medial branch blocks for chronic cervical facet joint pain: A randomized double-blind, controlled trial with one-year follow-up. *Spine (Phila Pa 1976)* 2008; 33:1813-1820.
 215. Manchikanti L, Singh V, Falco FJ, Cash KA, Pampati V. Lumbar facet joint nerve blocks in managing chronic facet joint pain: One-year follow-up of a randomized, double-blind controlled trial: Clinical Trial NCT00355914. *Pain Physician* 2008; 11:121-132.
 216. Manchikanti L, Manchikanti K, Manchukonda R, Cash KA, Damron KS, Pampati V, McManus CD. Evaluation of lumbar facet joint nerve blocks in the management of chronic low back pain: A preliminary report of a randomized, double-blind controlled trial: Clinical Trial NCT00355914. *Pain Physician* 2007; 10:425-440.
 217. Manchikanti L, Damron KS, Cash KA, Manchukonda R, Pampati V. Therapeutic cervical medial branch blocks in managing chronic neck pain: A preliminary report of a randomized, double-blind, controlled trial: Clinical Trial NCT0033272. *Pain Physician* 2006; 9:333-346.
 218. Agency for Healthcare Research and Quality, U.S. Department of Health and Human Services; Prepared by Southern California Evidence-based Practice Center. Assessment of the Need to Update Comparative Effectiveness Reviews: Report of an Initial Rapid Program Assessment (2005-2009). September 22, 2009.
 219. Clement FM, Harris A, Li JJ, Yong K, Lee KM, Manns BJ. Using effectiveness and cost-effectiveness to make drug coverage decision. A comparison of Britain, Australia, and Canada. *JAMA* 2009; 302:1437-1443.
 220. Koes B. Back pain research: Are patient registries ready for prime time? *Back Letter* 2009; 24:127-128.
 221. Deyo R. Back pain research: Are patient registries ready for prime time? *Back Letter* 2009; 24:127-128.
 222. Deyo RA, Mirza SK. The case for restraint in spinal surgery: Does quality management have a role to play? *Eur Spine J* 2008; 3:331-337.
 223. Chou R. Back pain research: Are patient registries ready for prime time? *Back Letter* 2009; 24:127-128.
 224. Tunis SR, Pearson SD. Coverage for promising technologies: Medicare's coverage with evidence development. *Health Affairs (Millwood)* 2006; 25:1218-1230.
 225. Malchau H, Herberts P, Eisler T, Garellick G, Söderman P. The Swedish Total Hip Replacement Register. *J Bone Joint Surg* 2002; 84:S2-S20.
 226. Eddy DM. Linking electronic medical records to large-scale simulation models: Can we put rapid learning on turbo? *Health Aff (Millwood)* 2007; 26:w137-w139.
 227. Berger ML, Mamdani M, Atkins D, Johnson ML. Good Research Practices for Comparative Effectiveness Research: Defining, Reporting and Interpreting Nonrandomized Studies of Treatment Effects Using Secondary Data Sources: The ISPOR Good Research Practices for Retrospective Database Analysis Task Force Report-Part I. *Value Health* 2009 Sep 29; [Epub ahead of print].
 228. Cox E, Martin BC, Van Staa T, Garbe E, Siebert U, Johnson ML. Good Research Practices for Comparative Effectiveness Research: Approaches to Mitigate Bias and Confounding in the Design of Nonrandomized Studies of Treatment Effects Using Secondary Data Sources: The International Society for Pharmacoeconomics and Outcomes Research Good Research Practices for Retrospective Database Analysis Task Force Report-Part II. *Value Health* 2009 Sept 10; [Epub ahead of print].
 229. Johnson, ML, Crown W, Martin BC, Dormuth CR, Siebert U. Good Research Practices for Comparative Effectiveness Research: Analytic Methods to Improve Causal Inference from Nonrandomized Studies of Treatment Effects Using Secondary Data Sources: The ISPOR Good Research Practices for Retrospective Database Analysis Task Force Report-Part III. 2009 Sep 29; [Epub ahead of print].