Comparative Effectiveness Topics from a Large, Integrated Delivery System

Kim N Danforth, ScD; Carrie D Patnode, PhD; Tanya J Kapka, MD; Melissa G Butler, PharmD, PhD; Bernadette Collins, PhD; Amy Compton-Phillips, MD; Raymond J Baxter, PhD; Jed Weissberg, MD, FACP; Elizabeth A McGlynn, PhD; Evelyn P Whitlock, MD

Perm J 2013 Fall;17(4):4-13

http://dx.doi.org/10.7812/TPP/13-036

Abstract

Objective: To identify high-priority comparative effectiveness questions directly relevant to care delivery in a large, US integrated health care system.

Methods: In 2010, a total of 792 clinical and operational leaders in Kaiser Permanente were sent an electronic survey requesting nominations of comparative effectiveness research questions; most recipients (83%) had direct clinical roles. Nominated questions were divided into 18 surveys of related topics that included 9 to 23 questions for prioritization. The next year, 648 recipients were electronically sent 1 of the 18 surveys to prioritize nominated questions. Surveys were assigned to recipients on the basis of their nominations or specialty. High-priority questions were identified by comparing the frequency a question was selected to an "expected" frequency, calculated to account for the varying number of questions and respondents across prioritization surveys. High-priority questions were those selected more frequently than expected.

Results: More than 320 research questions were nominated from 181 individuals. Questions most frequently addressed cardiovascular and peripheral vascular disease; obesity, diabetes, endocrinology, and metabolic disorders; or service delivery and systems-level questions. Ninety-five high-priority research questions were identified, encompassing a wide range of health questions that ranged from prevention and screening to treatment and quality of life. Many were complex questions from a systems perspective regarding how to deliver the best care.

Conclusions: The 95 questions identified and prioritized by leaders on the front lines of health care delivery may inform the national discussion regarding comparative effectiveness research. Additionally, our experience provides insight in engaging real-world stakeholders in setting a health care research agenda.

Introduction

Comparative effectiveness research has been proposed as a way to address the health care questions that are most relevant to patients, clinicians, and policymakers. Comparative effectiveness research is commonly defined as research designed to inform health care decision making through comparing the effectiveness, benefits, and harms of alternative strategies to diagnose, treat, or manage a clinical condition.[1] Currently, limited information exists regarding the effectiveness, benefits, and harms associated with many clinical practices. Furthermore, available research may not address the questions most relevant to practicing clinicians because studies may have included nonrepresentative patient groups in nonrepresentative settings (eg, academic medical centers), or have made comparisons to a placebo or untreated group. Comparative effectiveness research, in contrast, compares different strategies for preventing, diagnosing, treating, or managing a clinical condition in real-world settings with respect to their effectiveness, benefits, or harms. This type of research further seeks to determine what works best for whom, recognizing potential treatment response heterogeneity among populations. Thus, as the number of treatment and prevention options increases, and as appreciation of potential differences among individuals and populations grows, comparative effectiveness research has emerged as one way to improve the quality, efficiency, and value in health care.[2]

For comparative effectiveness research to reach its potential in improving and transforming health care, efforts will need to focus on the questions of greatest relevance to patients, clinicians, administrators, and policymakers. Integrated health care delivery organizations are well situated to identify important research questions whose answers could improve the everyday delivery of health care.[3] These settings include large groups of nonresearch and research clinicians who care for patients in a population-based model of care, health system administrators who manage the health care systems in which these patients are seen, and the members or patients themselves. Given the breadth of questions that comparative effectiveness research can address, identifying and prioritizing questions with the greatest clinical significance is essential and should include the perspective of practicing leaders and clinicians.

In 2009, Congress directed the Institute of Medicine (IOM) to identify national priorities for comparative effectiveness research to inform funding decisions by government agencies awarding grants under the American Recovery and Reinvestment Act.[4] When putting together its

Kim N Danforth, ScD, is a Research Scientist for the Department of Research and Evaluation, Kaiser Permanente Southern California in Pasadena. E-mail: kim.n.danforth@kp.org. **Carrie D Patnode, PhD,** is a Research Associate for the Center for Health Research-Northwest, Kaiser Permanente Northwest in Portland, OR. E-mail: carrie.d.patnode@kpchr.org. **Tanya J Kapka, MD,** is a Research Associate for the Center for Health Research-Northwest, Kaiser Permanente Northwest in Portland, and Lead Clinician at Virginia Garcia Memorial Health Center in Hillsboro, OR. E-mail: tkapka@vgmhc.org. **Melissa G Butler, PharmD, PhD,** is an Assistant Investigator for the Center for Health Research-Southeast, Kaiser Permanente Georgia in Atlanta. E-mail: melissa.g.butler@kp.org. **Bernadette Collins, PhD,** is a Consulting Practice Leader for the Division of Research in Oakland, CA. E-mail: bernie.collins@kp.org. **Amy Compton-Phillips, MD,** is the Associate Executive Director for The Permanente Federation in Oakland, CA. E-mail: amy.l.compton-phillips@kp.org. **Raymond J Baxter, PhD,** is the Senior Vice President of Community Benefit, Research and Health Policy for Kaiser Foundation Health Plan and Hospitals in Oakland, CA. E-mail: raymond.j.baxter@kp.org. **Jed Weissberg, MD, FACP,** is the Senior Vice President, Hospitals, Quality and Care Delivery Excellence for Kaiser Foundation Health Plan and Hospitals in Oakland, CA. E-mail: jed.weissberg@kp.org. **Elizabeth A McGlynn, PhD,** is the Director for the Kaiser Permanente Center for Effectiveness and Safety Research in Pasadena, CA. E-mail: elizabeth.a.mcglynn@kp.org. **Evelyn P Whitlock, MD,** is a Senior Investigator for the Center for Health Research-Northwest, Kaiser Permanente Northwest in Portland, OR. E-mail: evelyn.whitlock@kpchr.org.

list of the top 100 questions in comparative effectiveness, the IOM obtained input from a diverse group of stakeholders via interactive, mailed, and online mechanisms.[1] Nominated topics were reviewed and prioritized by the IOM committee, with additional topics added by the committee to diversify the portfolio.[1,5] Many questions on the IOM's list involved issues concerning health care delivery systems, racial and ethnic disparities, functional limitations and disabilities, and cardiovascular and peripheral vascular disease.

There have been several other efforts that engaged practicing clinicians or patients in identifying and prioritizing health care research questions, but most do not publish the actual questions prioritized. Instead, these efforts have focused on describing the methods of generation and prioritization of research questions.[6] As a result, it is unclear how well the questions that have been published or otherwise made widely available reflect the views of those on the front lines of health care delivery, who are key stakeholders and anticipated consumers of comparative effectiveness research.

In its report to Congress, the IOM recommended a "continuous evaluation of research topic priorities."[1] We conducted a survey of clinical and operational leaders within Kaiser Permanente (KP) to obtain their input on the comparative effectiveness research questions of particular importance to them. KP serves approximately nine million patients across the country and has been cited as one example of a large, preventive health care delivery system in national health care discussions. Thus, questions of high priority to KP leaders on the front lines of care delivery and health care decision making may be relevant to others. Additionally, it has been advocated that the questions generated by these types of surveys be published so that they are available to others.[6]

The aim of this article is to report the high-priority comparative effectiveness research questions identified and prioritized by practicing clinical and operational leaders in a large, diverse, integrated delivery system—along with the process used to engage them—to inform the national discussions on comparative effectiveness research.

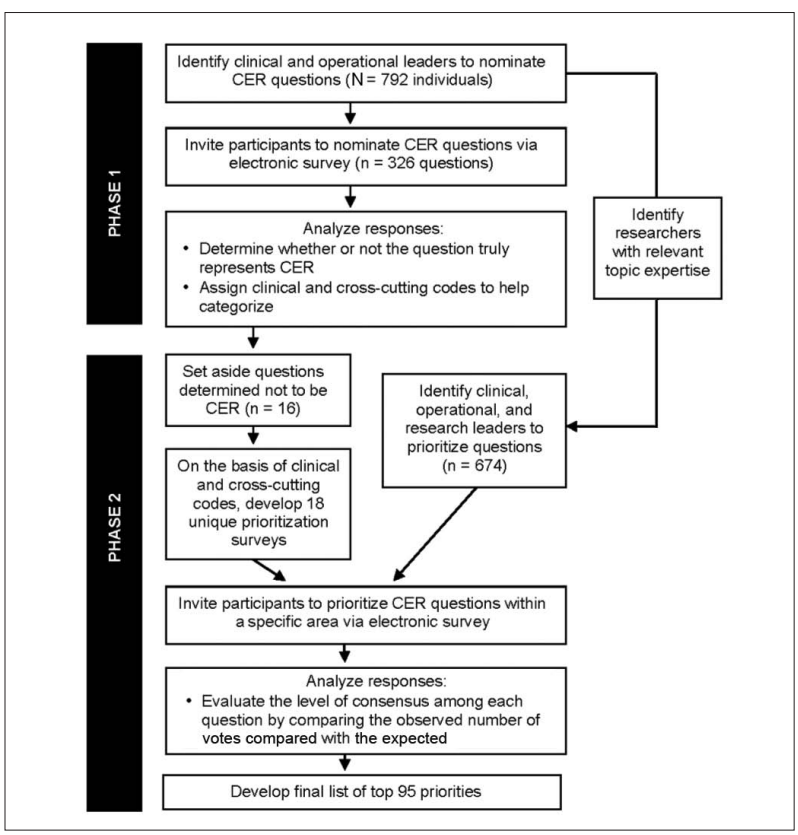

Figure 1. Project overview.

CER = comparative effectiveness research.

Methods

Study Setting

The KP Center for Effectiveness and Safety Research was established to promote and facilitate interregional research on effectiveness and safety involving the 8 KP Regions: Colorado, Georgia, Hawaii, Mid-Atlantic States (District of Columbia, Maryland, Virginia), Northern California, Northwest (Oregon and Washington), Ohio, and Southern California. KP is an integrated health care organization that provides comprehensive services to its members, including preventive, primary care, specialty, emergency, and hospital services. More than 15,000 physicians are employed by KP, and together the 8 Regions serve about 9 million members with diverse geographic, racial/ethnic, and socioeconomic characteristics. The work presented here was conducted as part of KP operational activities and was determined not to be research by the institutional review board.

Surveys of Clinical and Operational Leaders

To elicit comparative effectiveness research questions and subsequently prioritize them, we sent 2 surveys approximately 10 months apart to approximately 800 clinical and operational leaders in KP who were identified through input from national and regional executive leadership. Figure 1 displays the flow of identification and prioritization process.

Nomination of Questions

In Fall 2010, we e-mailed a link to the KP Survey of Critical Topics in Comparative Effectiveness to 792 clinical and operational leaders asking them to nominate up to 5 comparative effectiveness research questions within their areas of expertise. Recipients were invited to provide the specific comparative effectiveness research question and any relevant background information, including specific populations, interventions, outcomes, and comparators of interest. Survey recipients were identified in multiple ways, including

using existing distribution lists of clinical leaders in and across KP Regions for specific clinical specialties (eg, breast cancer, urology, cardiovascular disease, behavioral health), as well as lists of those involved in developing national KP clinical practice guidelines, working in quality improvement, or working in technology and products. Most survey recipients had direct clinical roles (83%); the remaining represented nonclinical roles such as executive leaders, experts in medical technology, and those in fields such as laboratory medicine.

Each nominated question was reviewed by the research team, and questions that were clearly not comparative effectiveness were excluded. For example, questions focused on establishing disease registries or clinical guidelines without mention of a specific comparative effectiveness research question were excluded. In making this determination, we used a broad definition guided by the IOM's definition of comparative effectiveness research.[1] Questions with multiple but distinct parts were separated (eg, if one part focused on prevention and another on treatment of a disease, they were separated into two questions). Likewise, nearly identical questions posed by different nominators were combined into a single research question.

Each question was reviewed by 2 team members to classify the question according to its content area from a listing of 43 possible codes. We assigned up to 3 clinical and 4 cross-cutting nonclinical (eg, pharmacology, service delivery) categories to each question. Clinical categories were adapted from the list used by the IOM[7] and were modified after pilot testing. Cross-cutting themes largely reflected overarching interests of the delivery organization, concerns among health care reformers, and clinical issues that did not fit into the more focused clinical conditions. Because the research team was particularly interested in questions related to cost, cost-effectiveness, and resource allocation, any question that contained this domain, either specifically in the question or in the background information provided by the nominator, was coded in this category. Additionally, a "main" classification was selected for each question from one of the clinical or cross-cutting classifications, favoring clinical areas unless the question primarily focused on a cross-cutting issue. Differences in classifications were resolved through informal discussion or team meetings. A final review of all questions was done by 1 team member (TJK) to ensure consistency of coding decisions across questions.

Table 1. Prioritization survey groups, number of questions, and recipients

Survey no.	Survey title	Number of questions on each survey	Number of stakeholders sent each survey
1	Chronic Diseases and Chronic Disease Management (includes Renal Diseases)	22	42
2	Cardiovascular Disease (Miscellaneous Topics)	15	36
3	Cardiovascular Disease (Pharmacy Topics)	18	36
4	Cardiovascular Disease (Surgery and Device Topics)	15	34
5	Ears, Nose, and Throat Disorders and Ophthalmology	9	26
6	Gastrointestinal System Disorders	9	29
7	Geriatrics	11	36
8	Health Systems	23	46
9	Infectious Diseases and Respiratory Diseases	17	36
10	Mental Health/Psychiatric Disorders and Addiction Medicine	15	40
11	Obesity, Diabetes, Endocrinology, and Metabolic Disorders	21	36
12	Oncology and Hematology	17	32
13	Pain Management	19	36
14	Palliative and End-of-Life Care/Dementia	14	34
15	Pediatrics	19	39
16	Prevention, Health Promotion, and Screening	13	38
17	Surgery, Procedures, and Anesthesia	21	38
18	Women's Health	10	34
Total		288	648[a]

[a] Twenty-six additional individuals were sent an invitation to complete a survey; however, the e-mail invitation was redirected back to us as undeliverable, likely indicating that they were no longer with the organization.

Prioritization of Questions

In the second phase of the project, we took the comparative effectiveness research questions generated by the KP comparative effectiveness research survey and further engaged KP stakeholders to prioritize among questions in broad clinical and systems-level categories. After omitting the research questions that were not clearly comparative effectiveness research and combining and splitting the questions as appropriate, a total of 288 questions remained. To facilitate prioritization, we divided the 288 questions into 18 groups of related topics (eg, obesity and diabetes). We believed that splitting the questions into smaller, more manageable lists of related topics would better facilitate prioritization rather than prioritizing across the full list of 288 questions. On the basis of this process, 18 electronic prioritization surveys were developed that included a range of 9 to 23 nominated research questions each (Table 1). The cardiovascular disease questions (n = 48) were separated into 3 surveys to make them more manageable for prioritization. In contrast, certain content areas received few nominations, and we elected to create more heterogeneous prioritization surveys containing these questions.

Because 10 months elapsed between nomination and prioritization, we updated the respondent list with input from national and regional KP leadership, including adding researchers with relevant expertise to the survey recipients. All of the original nominators and a random sample of the remaining group of original recipients were included in the updated list. The resulting 648 individuals were assigned to receive a particular prioritization survey based on their specialty area or whether they had nominated a question on that survey. Generally, recipients were sent only 1 prioritization survey, but there were a few exceptions (eg, someone nominated multiple questions that ended up on different prioritization surveys).

The number of individuals sent a particular prioritization survey ranged from 26 to 46.

Within each survey, we asked participants to choose the 5 research questions that they believed should have the highest priority for comparative effectiveness research in their set of grouped topics (without ranking them). Because the number of nominated questions and the number of survey respondents varied substantially across surveys, we sought a metric that would standardize the meaning of "highly prioritized" across content areas. Therefore, to identify highly prioritized questions, we calculated an "expected" number of times a question would be selected or "voted for" if all the questions were judged to be of equal importance and each reviewer selected 5 questions (no. of survey respondents multiplied by 5 potential votes/total no. of questions on the survey). For example, the pediatrics survey had 19 survey questions and 16 respondents, making the expected votes per question if all questions were judged to be of equal importance as follows: (16·5)/19 = 4.2. We then compared the number of times a question was actually selected (or "voted" for) with the expected number of votes to identify high-priority topics.

The *a priori* goal as stated at the outset by leaders initiating the survey was to generate a list of approximately 100 highly prioritized questions, which equated to approximately one-third of questions from each of the 18 surveys.

Table 2. First tier of high-priority comparative effectiveness research questions from surveys of clinical and operational leaders in Kaiser Permanente (n = 12)[a]

Survey topic area	Question/Objective
Chronic diseases and chronic disease management	Clinical effectiveness of alternatives to the physician's office visit in managing patients with hypertension and diabetes. Looking at quality metrics in these areas, and comparing performance on the basis of number of office visits per year, and the number and quality of telephone calls, vs HealthConnect messages. What is the most effective methodology, and does it vary by patient age?
	Compare the effectiveness (including resource utilization, workforce needs, net health care expenditures, and requirements for large-scale deployment) of new remote patient monitoring and management technologies (eg, telemedicine, Internet, remote sensing) and usual care in managing chronic disease.
	Would patients with chronic conditions (ie, diabetes, coronary artery disease, heart failure) receive more effective and efficient care delivery when care is delivered in a chronic care clinic when compared with traditional primary care clinic?
Geriatrics	What is the comparative effectiveness of different staffing models in the care of frail/functionally disabled seniors? 1. Traditional primary care physician-led primary care. 2. Geriatric care team with strong emphasis on registered nurse support. 3. Geriatric care team with strong emphasis on nurse practitioner support.
Health systems	Compare strategies to promote shared decision making by patients considering preference sensitive interventions vs usual care on decision outcome (screening choice, treatment choice, knowledge, treatment-preference concordance, and decisional conflict) with regard to various chronic conditions, including cancers. Possible strategies to compare include benefit designs and incentives; physician recommendation and "prescription"; and use as a "medical management" technique promoted by a health plan, payer, or employer.[b]
	Does colocation of behavioral health specialists (ie, social worker with a master of social work, licensed clinical social worker, psychologist, psychiatrist) in an adult primary care setting lead to improved patient outcomes including depression, anxiety, physical symptoms, physical disability, disease remission/modification (diabetes, hypertension, obesity, cardiovascular disease), quality of life, decreased absenteeism, prescription substance use, mental and physical function, satisfaction with the provider, and cost. Costs of care may include utilization of Emergency Department services, outpatient services, specialty psychiatry services, and total outpatient clinic visits.[b]
Mental health, psychiatry, and addiction medicine	Compare the effectiveness of case management approaches to standard individual psychotherapy in psychiatric patients with intermittent but persistent affective or anxiety disorders. Case management would include more flexible visits (eg, more frequent but shorter visit frequency based on symptom acuity, group visits, telephone visits, novel visits such as Internet chats).
Obesity, diabetes, endocrinology, and metabolic disorders	Compare the effectiveness of use of team-based approach of registered nurse (or other nonphysician) linked to a specific physician and his/her panel, with time for 3 to 20 contacts, to improve a HbA_{1C} > 9%, compared with usual care.
Oncology and hematology	Compare the effectiveness of management strategies for localized prostate cancer: active surveillance, androgen ablation, external beam radiotherapy, brachytherapy, radical retropubic prostatectomy, laparoscopic prostatectomy, robotic-assisted laparoscopic prostatectomy. Outcomes include quality of life, survival, recurrence, side effects, and cost.[b]
Pain management	Compare the effectiveness of treatment of chronic pain by cognitive behavioral and physical therapy treatment programs vs primary care treatment in an adult population.[b]
Pediatrics	Compare the effectiveness of the most common treatment modalities applied to symptoms associated with autism spectrum disorders, including physical therapy, occupational therapy, speech therapy, and applied behavior analysis-based strategies: stratified by the number of hours per week each is applied. The outcome would be measured on the basis of "effectiveness per hour" by modality, as measured by comparative pre- and postfunctional assessment by population.[b]
Surgery, procedures, anesthesia, and imaging	Compare the efficacy and morbidity of radical retropubic prostatectomy vs robotic prostatectomy for localized prostate cancer.

[a] Questions are listed in alphabetical order by survey topic area. Questions are listed as nominated, except for slight editing for style and combining questions on the same topic or separating out multipart questions on different topics.

[b] Survey question was a combination of separately nominated but nearly identical questions.

Table 3. Second tier of high-priority comparative effectiveness research questions from surveys of clinical and operational leaders in Kaiser Permanente (n = 46)[a]

Survey Topic Area	Question/Objectives
Cardiovascular disease: miscellaneous topics	Effectiveness of modeling: compare the decreased heart attacks and strokes from cardiovascular disease prevention program using the best results in the literature vs results from using a medical economic program (Archimedes) [software capable of modeling human physiology, diseases, behaviors, interventions, and health care systems] with direction to introduce principles of efficiencies and effectiveness to optimize outcome at controlled cost. Observational results from Archimedes are available; a comparable observational result would need to be chosen.
	Compare the effectiveness of advanced virtual care technologies for the home-based treatment of congestive heart failure vs conventional physical visits plus unilateral phone-based follow-up, for reducing unscheduled Emergency Department visits and hospitalizations in patients with severe disease.
Cardiovascular disease: pharmacy	Compare effectiveness of treating to target low-density lipoprotein cholesterol vs application of high-dose/high-potency statins for preventing cardiovascular events and mortality in persons with atherosclerotic cardiovascular disease or diabetes. Compare the effectiveness of using atorvastatin (Lipitor), 80 mg, vs rosuvastatin (Crestor), 40 mg (nonformulary), for patients who have coronary artery disease.[b]
	Compare the effectiveness of a "bundle of medications" for cardiovascular disease prevention with "usual care" of titration of blood pressure and lipids to target in individuals with high risk of cardiovascular disease.
	Compare the effectiveness of warfarin vs dabigatran (and other future oral anticoagulants) in 1) stroke prevention in atrial fibrillation patients, 2) intracranial hemorrhage, 3) gastrointestinal hemorrhage, 4) deep vein thrombosis or pulmonary embolism, 5) myocardial infarction rates, 6) gastrointestinal symptoms, 7) medication adherence, and 8) orthopedic prophylaxis.
Cardiovascular disease: surgery/ device/stent	Compare the effectiveness of treatment strategies for asymptomatic carotid stenosis including optimal medical management, carotid endarterectomy, and carotid artery stenting.
	Define the precise clinical settings in which percutaneous transluminal coronary angioplasty/stenting is superior to aggressive medical therapy for stable coronary disease. Need to include patients of different ages and with different comorbidities.[b]
	Compare the benefit of coronary artery bypass grafting vs aggressive medical therapy for coronary artery disease in patients who are being referred to surgery for control of angina.
	Compare the effectiveness of strategies for managing peripheral arterial disease in a patient population including screening programs, surveillance and tracking systems for monitoring known disease, and surgical and medical management options for treating advanced disease.
Chronic diseases and chronic disease management	Compare how chronic disease care gaps are resolved using the current model of health care delivery (primary care physician using an office face-to-face visit to address a care gap such as ordering a routine dual-energy x-ray absorptiometry, or DEXA, for primary screening) vs a systems approach where the DEXA request is generated from a list of all patients who need a DEXA and the DEXA order is prioritized on the basis of risk of hip and other fragility fractures. You could easily substitute any care gap such as Pap smear [Papanicolaou test], mammogram, or immunizations. The comparison is on face-to-face visits vs a systems approach that automatically requests the appropriate test.
	Address adherence to medications using a systematic approach to cover not just traditional barriers, but also mental/behavioral, cultural, literacy, and other personal issues. Identify those with adherence issues around crucial medications (PHASE [Preventing Heart Attacks and Strokes Everyday[c]] medications, for example, that impact outcomes) in that population, take a systematic patient-centered approach addressing all of the myriad of issues with tested interventions, and look at the impact on adherence and outcomes. Compare with similar population where adherence is identified and not addressed in a systematic way.
	Compare clinical outcomes (rate of complications, admissions to the hospital, mortality) of different models following stable in-center dialysis-dependent patients. More than once a month, once a month (currently required by Centers for Medicare and Medicaid Services), or less frequently, depending on clinical need. Use of physician extenders for in-center assessment vs a nephrologist's visit.
Ear, nose, and throat and ophthalmology	What is the most reliable method for calculating intraocular lens power following LASIK [laser-assisted in situ keratomileusis] in patients?
Gastrointestinal disorders	What is the comparative effectiveness of fecal immunochemical test alone annually, fecal immunochemical test plus flexible sigmoidoscopy every 5 years, or colonoscopy every 10 years with regard to death of colon cancer? Consider using a data-based national registry of all Kaiser Permanente patients screened and comparing outcomes.[b]
Geriatrics	What is the comparative effectiveness of health care provider home visits for the homebound frail/functionally disabled and the palliative care patient when compared with traditional clinic-based care?
	Compare effectiveness of providing geriatric primary care for frail/end-of-life members ages 85 years and older vs usual primary care.
Health systems	Compare satisfaction scores for patients seen in a clinic setting vs virtual visits.
	Among patients with mental health diagnoses seen in primary care practices, what is the effectiveness of components of integrated care services when compared with each other or with programs incorporating multiple components of integrated care, in leading to improved mental health outcomes?
	Compare the effectiveness of different benefit design, utilization management, and cost-sharing strategies in improving health care access and quality in patients with chronic conditions.
	Compare the use of systematic care coordination for a population who is high risk for utilization and cost vs a population of similar risk without systematic care coordination. Use the special needs plan model of care implemented in California, Colorado, and Georgia, and compare with nonspecial needs plan Regions with similar members to see any differences in outcomes—utilization, cost, and quality.
	Compare the effectiveness of dissemination and translation techniques to facilitate the use of comparative effectiveness research by patients, clinicians, payers, and others. How do we foster the adoption of proven strategies for diagnosis, treatment, and care into practice? What methodologies (electronic medical records, alerts, phone calls, incentives, etc) for translating findings into practice work the best and foster improved health outcomes at lower cost?

(continued on next page)

(continued from previous page)

Infectious diseases and respiratory disorders	Compare the effectiveness of preventing inpatient infections by performing daily chlorhexidine bed baths vs standard bathing.
	Compare the effectiveness of two weeks vs four weeks of intravenous antibiotics for treatment of *Staphylococcus aureus* bacteremia in patients with an identified transient source such as a catheter, abscess, or cellulitis.
	Compare the effectiveness of a noninvasive approach (serial lactate levels) to early goal-directed therapy for sepsis care with the current invasive approach. Study outcomes to be assessed would include overall mortality rates and costs of care (equipment and nursing care).[b]
	Compare the effectiveness of inhaled corticosteroids plus adherence encouragement techniques compared with inhaled corticosteroids plus long-acting β-agonists for treatment of persistent asthma uncontrolled on a regimen of medium-dose inhaled corticosteroids alone.
Mental health, psychiatry, and addiction medicine	Compare the effectiveness of inpatient drug and alcohol residential treatment vs outpatient chemical dependency program for substance abuse in long-term sobriety from drugs and alcohol.
Obesity, diabetes, endocrinology, and metabolic disorders	Does tight glycemic control improve outcomes and reduce cost of care of hospitalized patients?
	Compare the effectiveness of insulin pen devices vs insulin vial and syringe in type 2 diabetes.
	Compare the effectiveness and cost-effectiveness of conventional medical management of type 2 diabetes in adolescents and adults vs conventional therapy plus intensive educational programs or programs incorporating support groups and educational resources.
	What strategies are effective in getting large populations to eat a high-nutrition, low-calorie diet? Compare the effectiveness of financial incentives vs targeted advertising vs no intervention, to get patients to eat high-nutrition, low-calorie diets to treat obesity.
	Compare the effectiveness of treatments of metabolic syndrome and obesity: Mediterranean diet and exercise program with weekly group exercise sessions vs daily online interactive support/encouragement regarding diet/exercise (with or without quarterly group meetings in person).
	Comparative effectiveness and safety of bariatric surgery vs medications such as oral and/or injectable in the management of diabetic patients with HbA_{1C} above 8% and with body mass index over 35 kg/m^2.
Oncology and hematology	Compare the effectiveness in treatment outcomes in patients with a specific type of cancer that compares patients in whom a genetic biomarker is used to influence treatment options vs patients in whom the genetic biomarker is not useful.
Pain management	Compare the effectiveness of patients with chronic pain participating in a patient education/therapy program vs usual care through appropriate specialists and interventions without this additional training program.
	Compare the effectiveness of different nonnarcotic therapies used in the management of chronic pain/fibromyalgia.
	Evaluate the effectiveness of a coordinated pain policy in decreasing "drug-seeking behavior" in the Emergency Department and patient satisfaction.
Palliative and end-of-life care	Alzheimer disease/dementia: compare the effectiveness of screening and diagnostic strategies, as well as pharmacologic and nonpharmacologic treatments in treating the cognitive decline and slowing the functional decline of people with Alzheimer disease and other dementias, and managing behavioral disorders in home and institutional settings.[b]
	Comparative effectiveness of medical and nonmedical interventions in the behavioral symptoms of dementia.
	What treatments make a real difference in the quality of life for patients with Alzheimer disease?
	Compare the effectiveness of a palliative care consult on all patients living in long-term-care facilities vs no consult for decreasing hospital admissions, decreasing odds of dying in a hospital, decreasing overall cost of care in the last two months of life, and patient and family satisfaction.
Pediatrics	Compare the effectiveness of pediatric health screening questionnaires with physician conversation, health education classes, online handouts, online videos, and online interactive courses in improving outcomes for key health behaviors and risks such as sexually transmitted infections/pregnancy prevention, depression, firearm safety, and obesity.
	Compare the effectiveness of programs to prevent or to treat child obesity, as differentiated by degree of participation required and target areas of behavior change (reduce sugar-sweetened beverages, increase physical activity, decrease screen time, increase consumption of fruits and vegetables), on behavior change.
	What, if any, treatment makes a difference in the outcomes of children with autism spectrum disorder?
Prevention, screening, and health promotion	It has been said that there is no survival benefit for a population screened for prostate cancer (prostate specific antigen and digital rectal examination) compared with an "unscreened" population. The benefit in terms of reducing morbidity via screening is controversial. Screening for prostate cancer does result in a significant stage shift. Does this stage shift reflect a true reduction in morbidity, or is it simply time-lag bias?
	Compare the effectiveness of yearly health visit/interaction with a clinician on the quality of life of a patient vs no visit or interaction at all.
Surgery, procedures, anesthesia, and imaging	Study the impact of shared decision making in Kaiser Permanente on elective surgeries in a systematic approach over a large population to utilization, cost, risk management, and functional status and quality of life outcomes. Use a systematic intervention around shared decision making for a large group of members vs usual surgical consent and care in Kaiser Permanente. Although this has been studied in selected areas and some companies have implemented this approach (one example is Health Dialog, a care management/ decision support system [in Boston, MA]), this has not been rigorously studied in our own system in large populations as yet so we might see or demonstrate the impact. The hypothesis is that such an intervention in our system would reliably lead to even lower surgical rates, fewer bad outcomes, and less cost in risk management for a large number of elective surgeries, resulting in increased capacity, fewer direct and indirect costs, and better outcomes for members in terms of functional status and quality of life.

[a] Questions are listed in alphabetical order by survey topic area. Questions are listed as nominated, except for slight editing for style and combining questions on the same topic or separating out multipart questions on different topics.

[b] Survey question was a combination of separately nominated but nearly identical questions.

[c] PHASE is a cardiovascular risk reduction program that uses low-cost and generic medications and clinical interventions to reduce heart attacks, implemented at a systems level.

DEXA = dual energy x-ray absorptiometry.

We compared the observed number of votes with the expected number for each survey and categorized the questions into 3 tiers on the basis of the ratio of observed-to-expected votes; a cut-point of greater than the number of expected votes was used to identify high-priority questions because it yielded about 33% of nominated questions. The first tier consisted of questions receiving 2 or more times as many votes as expected; the second tier, those receiving 1.5 to 2 times the number of expected votes; and the third, those receiving more than the expected number but less than 1.5 times the expected number of votes.

Results

Nomination of Questions

Of the 792 individuals invited to nominate topics, 181 responded with at least 1 topic nomination (23% response rate). Most individuals (56%) nominated 1 topic, but 21% nominated 2 topics, and 23% nominated 3 or more. The nominators represented 50 distinct clinical specialties or areas of health system leadership.

A total of 326 research questions were received; 16 were dropped from prioritization because they were not comparative effectiveness research questions. After separating out distinct questions in a multiple-question nomination or combining nearly identical topics into a single question, there were 288 research questions for prioritization (Table 1). Questions on cardiovascular and peripheral vascular disease were the most frequent (n = 48 questions), followed by health systems (n = 23); chronic diseases and chronic disease management (n = 22); obesity, diabetes, endocrinology, and metabolic disorders (n = 21); and surgery, procedures, anesthesia, and imaging (n = 21). When question content was examined using the "main" topic classification according to our team's rating, independent of prioritization survey, similar results were observed. However, prevention, health promotion, and screening (n = 19) questions also were identified as a common focus of questions.

Many of the high-priority topics raised complex, systems-level questions about how to deliver the best care.

Prioritization: 95 High-Priority Questions

The overall prioritization survey response rate was 31%, ranging from 11% for the geriatrics survey to 53% for the oncology and hematology survey. Ninety-five questions were identified as high-priority questions on the basis of a comparison of the observed-vs-expected number of votes (Tables 2 to 4). There were 12 questions in the top tier, 46 in the second tier, and 37 in the third tier.

Of the 12 research questions in the top priority tier, 9 were questions from a systems perspective about the way in which care is delivered. For instance, questions focused on the comparative effectiveness of face-to-face vs remote management of patients (including different types of remote management); care provided in specialty clinics vs primary care; and the use of different staffing models (eg, linking nurses to a specific physician or emphasizing the role of nurses in care provision). In contrast, the other 3 top-tier questions focused on what specific care was best in particular clinical instances: 2 questions focused on treatment and management of prostate cancer, and 1 on treatment of autism. The prostate cancer questions both focused on localized prostate cancer and were related but not identical. One question proposed comparing a wide range of management and treatment methods, whereas the other focused on comparisons between radical retropubic vs robotic prostatectomy. The high-priority research questions in the second- and third-priority tiers represented a mix of broad systems-level and specific clinical questions (eg, comparisons of 2 drugs for a particular clinical condition).

The most common clinical categories among the high-priority questions were cardiovascular and peripheral vascular disease (19%); obesity, diabetes, endocrinology, and metabolic disorders (14%); and oncology and hematology (14%). Frequent cross-cutting, nonclinical areas were service delivery and systems-level issues (40%); pharmacology/pharmacy (34%); and prevention, health promotion, and screening (22%). Health information technology, which tended to include questions related to the electronic medical record, was also mentioned somewhat frequently (14%). Additionally, issues related to cost or cost-effectiveness were coded as occurring in most (62%) of the 95 high-priority research question nominations.

Discussion

The nominated and high-priority questions identified in this study ranged from prevention and screening to treatment and quality of life, reflecting the broad spectrum of issues encountered by practicing clinicians and administrators in a large health system. Questions addressed common health conditions facing our nation, including cardiovascular disease, obesity, and cancer, as well as topics related to health disparities, such as health literacy. Many of the high-priority topics raised complex, systems-level questions about how to deliver the best care.

Half of the 12 top-priority topics identified by our survey were the same or largely similar to questions on the IOM list. Overall, results from the 95 high-priority questions identified in our survey echoed some common themes from the IOM report, including health systems, chronic disease management, behavioral health integration into primary care, optimal cardiovascular disease management strategies, and concerns about better management of patients with chronic pain. In fact, despite favoring clinical areas in our determination of the "main" focus of a question, service delivery and systems-level questions were still the second most common main topic area.

The IOM prioritized system-level questions highly as well, with topics about health care delivery systems being the most common primary or secondary topic among its top 100 comparative effectiveness research questions.[1] However, the systems-level questions identified by KP leaders tended to be somewhat broader than those raised in the IOM's report. For instance, the high-priority questions in KP raised questions about staffing models (eg, primary care vs specialty care) or how care is delivered (eg, remote medicine vs in-person visit). In contrast, questions in the top quartile of the IOM's list focused more on comparisons of specific strategies for particular conditions. These comparisons encompassed wide-ranging

options (eg, primary prevention vs clinical interventions), but tended to focus more on the content of care rather than broad approaches to delivering care.

These differences between the IOM and KP lists may reflect differences in who was involved in nominating and prioritizing questions. Nearly all our respondents had direct patient care roles and practiced in an integrated delivery system, in contrast to the IOM respondents. Additionally, the processes themselves were different. For instance, the IOM solicited nominations from a wide group of stakeholders, started with many more nominated questions, and then prioritized them through several rounds of voting,[7] whereas we used a single round of prioritization among clinical and operational leaders in KP. The IOM also deliberately included questions of key significance to vulnerable subpopulations as part of its prioritization process.[1] Although we did not specifically seek out those types of questions, several nominated questions fell into that category, such as questions on health literacy and intimate partner violence, one of which was included among the 95 high-priority questions. Thus, the questions generated here are complementary to those of the IOM and reinforce the importance of certain questions and areas, such as care coordination, delivery, and management.

We designed the survey process around practical considerations, including respondent burden, which limited the information we collected. Although a large number of clinicians were invited to participate in the study, we did not randomly sample all staff in direct care-delivery roles. Instead, those invited to participate were intended to represent clinical and administrative leadership roles. Thus, the sample may not be fully reflective of all caregivers in our system or in general. Additionally, survey respondents were not equally inclusive across different specialties and did not include some specialties (eg, dentists).

We also grouped questions into 18 separate prioritization surveys instead of asking recipients to review all nominated questions to decrease respondent burden. Even in a prioritization survey, we did not ask recipients to rank all questions but rather to select their top 5 questions. Despite trying to minimize the time required to complete the surveys through these methods, response rates were relatively low, and thus the

Table 4. Third tier of high-priority comparative effectiveness research questions from surveys of clinical and operational leaders in Kaiser Permanente (n = 37)[a]

Survey Topic Area	Question
Cardiovascular disease: miscellaneous topics	Compare the effectiveness of ultrafiltration vs intravenous diuretics for patients admitted with heart failure and volume overload (alternative: compare the effectiveness of intermittent ultrafiltration vs diuretics for outpatients with severe heart failure and volume overload).
	Compare the outcomes specified as death or cardiovascular event between patients presenting with chest pain and unchanged or normal electrocardiogram and normal troponin levels, and who are in low to intermediate pretest probability; treated either by admission, observation, and early stress test, or by being placed on a regimen of β-blockers, aspirin, statins, and as-needed nitrates and sent home for outpatient stress testing.[b]
	Compare results and cost-effectiveness of conventional risk stratification (via stress testing, nuclear, etc) vs immediate computed tomographic angiography in patients who present to the Emergency Department with chest pain.
	Compare the effectiveness of computerized, individualized risk, and benefit assessment vs usual care (guidelines-driven treatment without computer models) for preventing cardiovascular events and mortality among people with atherosclerotic cardiovascular disease or diabetes.
Cardiovascular disease: pharmacy	Compare the effectiveness of warfarin (consistently maintained at an international normalized ratio of 2.0-3.0) plus aspirin, 81 mg, in adult patients with known coronary artery disease (who have a history of coronary artery bypass graft, percutaneous coronary intervention, myocardial infarction, etc) and an indication for thromboembolic prophylaxis (ie, atrial fibrillation) vs warfarin alone. Study outcomes should include myocardial infarction (fatal and nonfatal), coronary revascularization, stroke, and mortality as well as safety outcomes of major and minor bleeding.[b]
	Compare effectiveness of one vs two years of dual antiplatelet therapy post coronary stenting with drug-eluting stents for preventing recurrent myocardial infarctions, death, bleeding, and readmission for angina.
	Determine the incremental benefit of adding lipid-lowering agents to high-dose/high-potency statins for preventing cardiovascular events and mortality in persons with atherosclerotic cardiovascular disease or diabetes.
Chronic diseases and chronic disease management	Compare the effectiveness of care programs using patient-reported outcomes with those programs without patient-reported outcomes.
Ear, nose, and throat and ophthalmology	Compare effectiveness of antibiotic prophylaxis regimens for routine cataract surgery to prevent the devastating complication of endophthalmitis.[b]
	Compare effectiveness of screening for diabetic retinopathy by traditional means (face-to-face provider visits) vs photos reviewed by provider vs computer-read photo screening.
Gastrointestinal disorders	What is the optimal follow-up period after an adenomatous polyp has been removed during a colonoscopy?
	What is the value of upper endoscopy in the evaluation of gastroesophageal reflux disease?
Health systems	Compare the effectiveness of specialty clinical consults using traditional visits vs telemedicine visits.
Infectious diseases and respiratory disorders	Compare the effectiveness of real-time identification of uncontrolled asthma status linked to real-time notification of uncontrolled status to patients and asthma specialists to usual outreach asthma care management.
	Compare the effectiveness of treatment of community-acquired pneumonia with macrolides vs without macrolides.
Mental health, psychiatry, and addiction medicine	Compare the effectiveness of strategies to implement consistent treatment protocols for common mental disorders.
	What, if any, treatment makes a difference in the outcomes of adults with autism spectrum disorder?
	Compare the effectiveness of usual care (a reactive behavioral model) vs a care management program (an integrated, proactive medical model that engages patients) for patients using the Emergency Department for alcohol/substance abuse-related problems.

(continued on next page)

nominated and prioritized topics may not be representative of KP clinicians and leaders as a whole. Low response rates were also a problem in the high-profile IOM process, in which about 9% of those e-mailed responded with nominations.[1] Overall, the limitations due to how we grouped the questions and conducted the study, along with the low response rates, are important to consider when assessing the internal and external generalizability of the results. Issues in particular disease areas may be underrepresented or overrepresented given the number of individuals invited to participate and responders in that area. However, we surveyed a wide range of clinical and operational leaders, whose responses correspondingly reflected the wide spectrum of issues faced by leaders in an integrated delivery system. Although different prioritization methods or a higher response rate (and therefore a different group of responders) may have yielded different high-priority questions, we believe that the questions identified here are still likely to represent questions of practical, clinical importance given that they were first nominated and then prioritized by a diverse group of clinicians and administrators.

Given the number of nominations we received, it was out of scope for this project to include more objective information regarding the underlying disease burden or to qualify nominated questions as clearly unanswered (by searching for in-process or published research). Thus, we cannot confirm whether the nominators' questions represented needed research or instead indicated lack of dissemination of existing research findings or recommendations for practice. In other work, we have found that about 25% of the time, publicly nominated research questions are already addressed through recent systematic reviews (Michelle Eder, PhD, oral communication, 2012 May 14).[a]

Another limitation is that we were not able to determine why highly prioritized questions were selected. We collected ratings from the nominators and prioritizers regarding the potential impact of the question on health care quality, efficiency, or equity. However, questions were generally rated highly on all these domains, which did not enable us to discriminate the reason for the priority. Additionally, we did not include patients in our surveys. However, their perspective is being obtained by the Patient-Centered Outcomes Research Institute, which is currently asking patients to nominate research questions. This research institute also is encouraging studies to include patients and other stakeholders in the research process by making their involvement part of the criteria for funding decisions, as well as including patients in the review of submitted proposals.[8]

(continued from previous page)

Oncology and hematology	Evaluate whether ductal carcinoma in situ needs treatment.
	Standard follow-up with or without positron emission tomography scan for follow-up for lymphoma, with death as endpoint. When is positron emission tomography needed?
	Compare hospice care vs usual medical care for patients with Stage 4 (advanced) cancer, with respect to survival, quality of life, and health care costs.
	Does a multidisciplinary breast cancer clinic improve care for patients with breast cancer compared with a "traditional" nonintegrated care system?
Pain management	Compare outcomes in acute lumbar radiculopathy treated with oral steroid vs no steroid.
Palliative and end-of-life care	Compare the effectiveness of a trigger-based vs a referral-based palliative care program.
	Does the presence of an inpatient palliative care service improve care for patients with cancer?
	Would a video-based tool describing end-of-life choices for dementia, made available via Internet or kiosk/checkout at physician's office compared with usual option of advance directive written brochure (Your Life/Your Choices class) increase the number of patients with an advance directive or (physician's orders for life-sustaining treatment) in patients with early dementia?
Pediatrics	How important is it that patients have a registered visit to assess their progress on attention deficit disorder/attention deficit with hyperactivity disorder medication at various intervals after beginning their medication? Would a documented phone call that may not be a telephone advice visit, in our medical record, be a legitimate option?
Prevention, screening, and health promotion	Compare the effectiveness of a program of incenting patients to stop smoking vs usual smoking cessation care, to test whether providing incentives to members who stop smoking is a more effective strategy.
	Compare the effectiveness of traditional behavioral interventions vs economic incentives in motivating behavior changes (eg, weight loss, smoking cessation, avoiding alcohol and substance abuse) in children and adults.
Surgery, procedures, anesthesia, and imaging	Compare the cost and outcomes of benign gynecology hysterectomy methods: vaginal, laparoscopic, abdominal.
	Are there any short- or long-term benefits/risks from the use of steroid injections or hylan G-F 20 (Synvisc) in treating knee arthritis?
	Compare the effectiveness of having surgery vs not having surgery for arthroscopy.
	Compare the effectiveness of the use of INFUSE bone morphogenic protein, a proprietary product from Medtronic, to alternative biologics/growth factors like allograft, human-derived demineralized bone matrix, and synthetic (ie, ceramic) bone graft substitutes.
Women's health	Conduct a prospective trial comparing elective cesarean delivery vs induced vaginal delivery for pregnant women who are 41 or more weeks' gestation with an unfavorable Bishop score and high fetal vertex presentation. Look at maternal and neonatal morbidity and overall length of stay.
	Compare the effectiveness of induction of labor using extraordinary levels of oxytocin over routine methods of induction for all matched pregnancies.
	Compare rates of thrombosis between different forms of hormone replacement therapy for menopause.
	Compare the effectiveness of screening methods for improving identification of intimate partner violence in female Health Plan members: routine universal screening vs screening of patients with intimate partner violence risk factors.

[a] Questions are listed in alphabetical order by survey topic area. Questions are listed as nominated, except for slight editing for style and combining questions on the same topic or separating out multipart questions on different topics.

[b] Survey question was a combination of separately nominated but nearly identical questions.

A range of quantitative and qualitative methods have been used to obtain stakeholders' input on research needs and priorities. These methods have included semistructured, rating, and forced-ranking surveys, iterative Delphi or modified Delphi techniques, focus groups, citizens' juries and consumer panels (in the case of engaging the public at-large), and deliberative democracy and other facilitated consensus-building methods.[9] Each of these methods has different strengths and weaknesses, and their most appropriate application depends on context (eg, the intended focus of the research, its intended uses, and the available resources to support the engagement). Our method was a broad-based initial effort to engage real-world health care leaders in a large population-based system that has an integrated research and quality-improvement capability, with the potential for further development. As other organizations deliberate processes for eliciting research needs and priorities, an essential step is for them to think about the ultimate implications and uses of the results. They may wish to draw on existing work, such as the Agency for Healthcare Research and Quality's prioritization of topics by reviewing existing evidence (eg, prevalence, mortality, variations in treatment, existing studies).[7,10]

Conclusion

In conclusion, by seeking input from practicing clinicians and operational leaders in a large health system providing comprehensive care, we obtained a wide range of questions reflecting the diverse health issues facing patients, clinicians, and health care systems. The 95 high-priority questions presented here represent issues of importance to those on the front lines of health care delivery, who are key stakeholders and anticipated consumers of comparative effectiveness research. Thus, these questions may help inform the national discussion regarding comparative effectiveness research and health care. ❖

[a] *Research Associate, Kaiser Permanente Center for Health Research, Portland, OR*

Disclosure Statement

This study was funded through Kaiser Permanente internal operating funds provided by the Kaiser Permanente Center for Effectiveness and Safety Research. The author(s) have no other conflicts of interest to disclose.

Acknowledgment

We gratefully acknowledge the contributions and support provided by: Joe Selby, MD, Executive Director, Patient-Centered Outcomes Research Institute, Washington, DC; the Kaiser Permanente National Research Council; and the Kaiser Foundation Research Institute, Oakland, CA.

Kathleen Louden, ELS, of Louden Health Communications provided editorial assistance.

References

1. Committee on Comparative Effectiveness Research Prioritization, Institute of Medicine of the National Academies. Initial national priorities for comparative effectiveness research: report brief [monograph on the Internet]. Washington, DC: Institute of Medicine of the National Academies; 2009 Jun 30 [cited 2013 Aug 14]. Available from: www.iom.edu/~/media/Files/Report%20Files/2009/ComparativeEffectivenessResearchPriorities/CER%20report%20brief%2008-13-09.pdf.
2. Federal Coordinating Council for Comparative Effectiveness Research. Report to the President and the Congress. Washington, DC: US Department of Health and Human Services; 2009 Jun 30.
3. Dubois RW, Graff JS. Setting priorities for comparative effectiveness research: from assessing public health benefits to being open with the public. Health Aff (Millwood) 2011 Dec;30(12):2235-42. DOI: http://dx.doi.org/10.1377/hlthaff.2011.0136
4. American Recovery and Reinvestment Act of 2009, Pub L No. 111-5 (Feb 17, 2009).
5. Iglehart JK. Prioritizing comparative-effectiveness research—IOM recommendations. N Engl J Med 2009 Jul 23;361(4):325-8. DOI: http://dx.doi.org/10.1056/NEJMp0904133
6. Stewart RJ, Caird J, Oliver K, Oliver S. Patients' and clinicians' research priorities. Health Expect 2011 Dec;14(4):439-48. DOI: http://dx.doi.org/10.1111/j.1369-7625.2010.00648.x
7. Committee on Comparative Effectiveness Research Prioritization, Institute of Medicine of the National Academies. Initial national priorities for comparative effectiveness research. Washington, DC: Institute of Medicine of the National Academies; 2009.
8. Patient-Centered Outcomes Research Institute [homepage on the Internet]. Washington, DC: Patient-Centered Outcomes Research Institute; c2013 [cited 2013 Aug 14]. Available from: www.pcori.org.
9. O'Haire C, McPheeters M, Nakamoto E, et al; Oregon Evidence-based Practice Center; Vanderbilt Evidence-based Practice Center. Engaging stakeholders to identify and prioritize future research needs. Methods future research needs report number 4. Publication No. 11-EHC044-EF [monograph on the Internet]. Rockville, MD: Agency for Healthcare Research and Quality; 2011 Jun [cited 2013 Aug 14]. Available from: http://effectivehealthcare.ahrq.gov/ehc/products/200/698/MFRNGuide04--Engaging_Stakeholders--6-10-2011.pdf.
10. Whitlock EP, Lopez SA, Chang S, Helfand M, Eder M, Floyd N. AHRQ series paper 3: identifying, selecting, and refining topics for comparative effectiveness systematic reviews: AHRQ and the effective health-care program. J Clin Epidemiol 2010 May;63(5):491-501. DOI: http://dx.doi.org/10.1016/j.jclinepi.2009.03.008

Joy

When, after so many efforts, you have at last arrived at a certainty,
your joy is one of the greatest which can be felt by a human soul.

— Louis Pasteur, 1822-1895, French chemist and microbiologist