Draft Comparative Effectiveness Review

Number XX

Communication and Dissemination Strategies to Facilitate the Use of Health Related Evidence

Prepared for:
Agency for Healthcare Research and Quality
U.S. Department of Health and Human Services
540 Gaither Road
Rockville, MD 20850
www.ahrq.gov

Contract No.

Prepared by:

Investigators:
First and Last Names, X.X.
First and Last Names, X.X.

AHRQ Publication No. xx-EHCxxx
<Month Year>

Statement of Funding and Purpose

This report is based on research conducted by the <NAME>Evidence-based Practice Center (EPC) under contract to the Agency for Healthcare Research and Quality (AHRQ), Rockville, MD (Contract No. <NUMBER>). The findings and conclusions in this document are those of the authors, who are responsible for its contents; the findings and conclusions do not necessarily represent the views of AHRQ. Therefore, no statement in this report should be construed as an official position of AHRQ or of the U.S. Department of Health and Human Services.

The information in this report is intended to help health care decisionmakers—patients and clinicians, health system leaders, and policymakers, among others—make well-informed decisions and thereby improve the quality of health care services. This report is not intended to be a substitute for the application of clinical judgment. Anyone who makes decisions concerning the provision of clinical care should consider this report in the same way as any medical reference and in conjunction with all other pertinent information, i.e., in the context of available resources and circumstances presented by individual patients.

This report may be used, in whole or in part, as the basis for development of clinical practice guidelines and other quality enhancement tools, or as a basis for reimbursement and coverage policies. AHRQ or U.S. Department of Health and Human Services endorsement of such derivative products may not be stated or implied.

Public Domain Notice

This document is in the public domain and may be used and reprinted without permission except those copyrighted materials that are clearly noted in the document. Further reproduction of those copyrighted materials is prohibited without the specific permission of copyright holders.

Disclaimer Regarding 508-Compliance

Persons using assistive technology may not be able to fully access information in this report. For assistance contact [insert program e-mail address].

Financial Disclosure Statement

None of the investigators has any affiliations or financial involvement that conflicts with the material presented in this report.

[Standard Suggested Citation of Your Report]

Suggested citation: <Authors>. Title. Comparative Effectiveness Review No. xx. (Prepared by the <NAME> Evidence-based Practice Center under Contract No. <NUMBER> .) AHRQ Publication No. xx-EHCxxx. Rockville, MD: Agency for Healthcare Research and Quality. <Month, Year>. http://www.ahrq.gov/clinic/epcix.htm. Accessed: <Date>. <URL>.

Preface

The Agency for Healthcare Research and Quality (AHRQ) conducts the Effective Health Care Program as part of its mission to organize knowledge and make it available to inform decisions about health care. As part of the Medicare Prescription Drug, Improvement, and Modernization Act of 2003, Congress directed AHRQ to conduct and support research on the comparative outcomes, clinical effectiveness, and appropriateness of pharmaceuticals, devices, and health care services to meet the needs of Medicare, Medicaid, and the Children's Health Insurance Program (CHIP).

AHRQ has an established network of Evidence-based Practice Centers (EPCs) that produce Evidence Reports/Technology Assessments to assist public- and private-sector organizations in their efforts to improve the quality of health care. The EPCs now lend their expertise to the Effective Health Care Program by conducting comparative effectiveness reviews (CERs) of medications, devices, and other relevant interventions, including strategies for how these items and services can best be organized, managed, and delivered.

Systematic reviews are the building blocks underlying evidence-based practice; they focus attention on the strength and limits of evidence from research studies about the effectiveness and safety of a clinical intervention. In the context of developing recommendations for practice, systematic reviews are useful because they define the strengths and limits of the evidence, clarifying whether assertions about the value of the intervention are based on strong evidence from clinical studies. For more information about systematic reviews, see

http://www.effectivehealthcare.ahrq.gov/reference/purpose.cfm

AHRQ expects that CERs will be helpful to health plans, providers, purchasers, government programs, and the health care system as a whole. In addition, AHRQ is committed to presenting information in different formats so that consumers who make decisions about their own and their family's health can benefit from the evidence.

Transparency and stakeholder input from are essential to the Effective Health Care Program. Please visit the Web site (http://www.effectivehealthcare.ahrq.gov) to see draft research questions and reports or to join an e-mail list to learn about new program products and opportunities for input. Comparative Effectiveness Reviews will be updated regularly.

We welcome comments on this CER. They may be sent by mail to the Task Order Officer named below at: Agency for Healthcare Research and Quality, 540 Gaither Road, Rockville, MD 20850, or by e-mail to epc@ahrq.hhs.gov.

Carolyn M. Clancy, M.D.
Director
Agency for Healthcare Research and Quality

Jean Slutsky, P.A., M.S.P.H.
Director, Center for Outcomes and Evidence
Agency for Healthcare Research and Quality

Marian James, Ph.D., M.A.
Task Order Officer
Evidence-based Practice Program
Center for Outcomes and Evidence
Agency for Healthcare Research and Quality

William Lawrence, M.D., M.S.
Task Order Officer
Evidence-based Practice Program
Center for Outcome and Evidence
Agency for Healthcare Research and Quality

Acknowledgments

The authors gratefully acknowledge the continuing support of our AHRQ Task Order Officer, Marian James, PhD, MA and William Lawrence, MD, MS. We extend our appreciation to our Key Informants and members of our Technical Expert Panel (listed below), all of whom provided thoughtful advice and input during our research process.

The investigators deeply appreciate the considerable support, commitment, and contributions of the EPC team staff at <NAME>. We express our gratitude to the following individuals for their contributions to this project: <NAME, degrees>

Key Informants

Key Informants are the end users of research, including patients and caregivers, practicing clinicians, relevant professional and consumer organizations, purchasers of health care, and others with experience in making health care decisions. Within the EPC program, the Key Informant role is to provide input into identifying the Key Questions for research that will inform health care decisions. The EPC solicits input from Key Informants when developing questions for systematic review or when identifying high priority research gaps and needed new research. Key Informants are not involved in analyzing the evidence or writing the report and have not reviewed the report, except as given the opportunity to do so through the peer or public review mechanism

Key Informants must disclose any financial conflicts of interest greater than $10,000 and any other relevant business or professional conflicts of interest. Because of their role as end-users, individuals are invited to serve as Key Informants and those who present with potential conflicts may be retained. The Task Order Officer and the EPC work to balance, manage, or mitigate any potential conflicts of interest identified.

Key Informants

To be completed after peer review.

Technical Expert Panel

Technical Experts comprise a multidisciplinary group of clinical, content, and methodologic experts who provide input in defining populations, interventions, comparisons, or outcomes as well as identifying particular studies or databases to search. They are selected to provide broad expertise and perspectives specific to the topic under development. Divergent and conflicted opinions are common and perceived as healthy scientific discourse that results in a thoughtful, relevant systematic review. Therefore study questions, design and/or methodological approaches do not necessarily represent the views of individual technical and content experts. Technical Experts provide information to the EPC to identify literature search strategies and recommend approaches to specific issues as requested by the EPC. Technical Experts do not do analysis of any kind nor contribute to the writing of the report and have not reviewed the report, except as given the opportunity to do so through the peer or public review mechanism.

Technical Experts must disclose any financial conflicts of interest greater than $10,000 and any other relevant business or professional conflicts of interest. Because of their unique clinical or content expertise, individuals are invited to serve as Technical Experts and those who present with potential conflicts may be retained. The Task Order Officer and the EPC work to balance, manage, or mitigate any potential conflicts of interest identified.

Technical Expert Panel

To be completed after peer review.

Peer Reviewers

Peer Reviewers are invited to provide written comments on the draft report based on their clinical, content, or methodologic expertise. Peer review comments on the preliminary draft of the report are considered by the EPC in preparation of the final draft of the report. Peer Reviewers do not participate in writing or editing of the final report or other products. The synthesis of the scientific literature presented in the final report does not necessarily represent the views of individual reviewers. The dispositions of the peer review comments are documented and will, for CERs and Technical briefs, be published three months after the publication of the Evidence report.

Potential Reviewers must disclose any financial conflicts of interest greater than $10,000 and any other relevant business or professional conflicts of interest. Invited Peer Reviewers may not have any financial conflict of interest greater than $10,000. Peer Reviewers who disclose potential business or professional conflicts of interest may submit comments on draft reports through the public comment mechanism.

Peer Reviewers

To be completed after peer review

Communication and Dissemination Strategies to Facilitate the Use of Health Related Evidence

Structured Abstract

Objectives. Accelerating awareness and adoption of health related evidence to inform health and health care decisions by patients and caregivers, practicing clinicians, relevant professional and consumer organizations, purchasers of health care, and others with experience in making health care decisions can be accomplished if effective communication and dissemination strategies are identified and applied. For decades, researchers and other stakeholders have explored approaches for communicating and disseminating evidence to close the research-to-practice gap. This review sought to provide evidence for how to best communicate and disseminate evidence, including uncertain evidence to inform health care decisions. The review focused on three primary objectives: (1) address the comparative effectiveness of communicating evidence in various contents and formats that increase the likelihood that target audiences will both understand and use the information; (2) examine the comparative effectiveness of a variety of approaches for disseminating evidence from those who develop it to those who are expected to use it; and (3) examine the comparative effectiveness of various ways of communicating uncertainty associated health-related evidence to different target audiences. A secondary objective was to examine how the effectiveness of communication and dissemination strategies varies across target audiences, namely patients and clinicians.

Data sources. We searched MEDLINE®, the Cochrane Library, Cochrane Central Trials Registry, PsychInfo, and the Web of Science. We used a variety of medical subject headings (MeSH terms) and major headings, free-text and title and abstract text-word searches, limited to studies on humans published from 2000 to May 15, 2012 for communication and dissemination given the prior systematic reviews, and from 1966 onward to May 15, 2012 for communicating uncertainty.

Review methods. We used standard Evidence-based Practice Center methods of dual review of abstracts, full-text articles, abstractions, and quality ratings and group consensus to resolve disagreements. We used group consensus to grade strength of evidence.

Results. The initial search identified 3,917 articles (after removing duplicates) for all three KQs. After dual review at the title/abstract stage and full text review stage, we retained 53 articles that directly (i.e., head to head) compared strategies to communicate and disseminate evidence. Six articles (representing five trials) looked at the effect of various communication strategies (KQ1); 40 articles (representing 37 trials) focused on evidence dissemination to clinicians or patients (KQ2); and eight articles (representing seven trials) examined alternate ways to communicate the overall strength of evidence, and the precision, directness, and net benefit of evidence (KQ3). For each KQ, a summary of overall findings, limitations, future directions, and implications for patients, clinicians, and policymakers is presented.

Conclusions. The lack of research evidence to inform communication and dissemination of evidence, including uncertain evidence impedes timely clinician, patient and policymaker awareness, uptake and use of evidence to improve the quality of care. Expanding investment in communication, dissemination and implementation research is critical to the identification of

strategies to accelerate the translation of comparative effectiveness research into community and clinical practice and the direct benefit of patient care.

More research is needed to better understand the current barriers to translating the findings of comparative effectiveness research into community and clinical practice. Ongoing funding for interdisciplinary communication and dissemination sciences research is needed to promote the uptake and use of evidence and ensure quality of care.

Contents

Executive Summary ES-1

Introduction 1
Background 1
Rationale and Relevance for Conducting the Systematic Review 1
Terminology and Definitions 1
Communication Strategies to Promote the Use of Health Care Evidence 2
Dissemination Strategies to Promote the Use of Health Care Evidence 5
Explaining Uncertain Evidence 6
Scope and Key Questions 8
Scope of the Review 8
Key Questions 8
Analytic Framework 9
Populations, Interventions, Comparators, Outcomes, Timeframes, and Settings Covered By the Key Questions 10
Organization of This Report 10
Methods 12
Topic Refinement and Review Protocol 12
Literature Search Strategy 12
Search Strategy 12
Inclusion and Exclusion Criteria: Overall 13
Inclusion and Exclusion Criteria Specific to Communication Techniques 16
Inclusion and Exclusion Criteria Specific to Dissemination Strategies 17
Inclusion and Exclusion Criteria for Studies to Present Uncertain Evidence 19
Study Selection 21
Data Extraction 21
Risk of Bias Assessment of Individual Studies 22
Data Synthesis: Overall 23
Data Synthesis: Methods Specific to Key Questions 23
Strength of the Body of Evidence 23
Applicability 24
Peer Review and Public Commentary 25
Results: Key Question 1: Communication Strategies 26
Results of Literature Searches for Key Questions 1, 2 and 3 26
Key Question 1: Communicating Evidence to Patients and Clinicians 27
Key Points 27
Detailed Synthesis 29
Results: Key Question 2: Dissemination Strategies to Clinicians and Patients 37
Introduction 37
Description of Included Studies 37
Key Question 2: Disseminating Evidence to Clinicians 38
Key Points: Clinician Trials 38
Detailed Synthesis: Dissemination Strategies for Clinicians and Health-related Decisions or Behaviors 39
Detailed Synthesis: Dissemination Strategies for Clinicians and Clinical Outcomes 48

Detailed Synthesis: Dissemination Strategies for Clinicians and Knowledge Outcomes 52
Key Question 2: Disseminating Evidence to Patients 56
Key Points: 56
Detailed Synthesis: Dissemination Strategies for Patients and Health-Related Decisions or Behaviors 56
Detailed Synthesis: Dissemination Strategies to Patients for Clinical Outcomes 64
Detailed Synthesis: Dissemination Strategies for Patients and Knowledge Outcomes 67
Key Question 2: Disseminating Information to Clinicians and Patients 71
Key Points 71
Detailed Synthesis: Dissemination Strategies for Clinicians and Patients for Health-related Decisions and Behavior Outcomes 71
Detailed Synthesis: Dissemination Strategies for Clinicians and Patients for Clinical Outcomes 76
Results: Key Question 3: Communicating Uncertainty 78
Introduction 78
Description of Included Studies 78
Key Question 3: The Effects of Communicating Uncertain Health and Health Care Evidence to Patients and Clinicians 79
Key Points 79
Detailed Synthesis 83
Discussion.....95
Key Question 1: Communication Strategies 95
Key Findings and Strength of Evidence 95
Findings in Relationship to What is Already Known 95
Mode of Communication 96
Applicability 97
Research Gaps 98
Key Question 2: Dissemination Strategies 99
Key Findings and Strength of Evidence 99
Findings in Relationship to What is Already Known 99
Limitations of the Evidence Base or the Review Procedures 100
Applicability 102
Research Gaps 102
Key Question 3: Communication of Uncertainty 103
Key Findings and Strength of Evidence 103
Findings in Relationship to What is Already Known 104
Limitations in the Evidence Base and Review Procedures 104
Applicability 104
Research Gaps 105
Evidence in the Broader Context.....105
Conceptual Issues 106
Methodological Considerations Regarding Our Review 108
Future Research 110
Implications for Patients, Clinicians, and Policymakers 110
Conclusions 111
References 112

Tables

Table 1. Definitions of concepts relevant for this review 2
Table 2. Systematic, meta-analytic, or theoretical reviews supporting various communication strategies 3
Table 3. Conceptual framework for uncertainty (KQ 3) 7
Table 4. Population, intervention, comparators, outcomes, and settings (PICOTS) 10
Table 5. Initial literature search terms for each of the targeted searches 13
Table 6. General inclusion/exclusion criteria for all key questions 14
Table 7. Inclusion and exclusion criteria for research evidence to be communicated or disseminated 15
Table 8. Included communication strategies and approaches for key question 1 16
Table 9. Included dissemination strategies for key question 2 18
Table 10. Included communication strategies and approaches for key question 3 20
Table 11. Data items extracted 22
Table 12. Definitions of the grades of overall strength of evidence 24
Table 13. Strength of evidence of communication strategies 28
Table 14. Studies of communication strategies 30
Table 15. Strength of evidence: dissemination to clinicians for health-related decisions and behavior outcomes 39
Table 16. Summary of trials examining dissemination to clinicians for health-related decisions and behavior outcomes 41
Table 17. Strength of evidence: Studies examining dissemination to clinicians for clinical outcomes 48
Table 18. Summary of trials examining dissemination to clinicians for clinical outcomes 49
Table 19. Strength of evidence: Trials examining dissemination to clinicians for knowledge outcomes. 53
Table 20. Summary of studies examining dissemination to clinicians for knowledge outcomes 54
Table 21. Strength of evidence: Trials examining dissemination to patients for health-related decisions and behavior outcomes. 57
Table 22. Summary of trials examining dissemination to patients for health-related decisions and behavior outcomes 58
Table 23. Strength of evidence: Trials examining dissemination to patients for clinical outcomes 64
Table 24. Summary of trails examining dissemination to patients for clinical outcomes 65
Table 25. Strength of evidence: Trails examining dissemination to patients for knowledge outcomes 67
Table 26. Summary of trails examining dissemination to patients for knowledge outcomes 68
Table 27. Strength of evidence: Trials examining dissemination to clinicians and patients for health-related decisions and behavior outcomes 72
Table 28. Summary of studies examining dissemination to clinicians and patients for health-related decisions and behavior outcomes 73
Table 29. Strength of evidence: Trial examining dissemination to clinicians and patients for clinical outcomes 76
Table 30. Summary of studies examining dissemination to clinicians and patients for clinical outcomes 77

Table 31. Strength of evidence: Communicating strength of evidence 80
Table 32. Strength of evidence: Communicating precision 81
Table 33. Strength of evidence: Communicating directness 81
Table 34. Strength of evidence: Communicating net benefit 82
Table 35. Studies of communicating strength of evidence 84
Table 36. Studies of communicating precision 86
Table 37. Study on directness 89
Table 38. Studies about providing information about net benefit (balance of benefits and harms) 90

Figures

Figure A. Analytic framework for communicating and disseminating evidence and explaining uncertainty ES-4
Figure B. Evidence continuum in implementation science ES-10

Figure 1. Analytic framework for communicating and disseminating strategies and explaining uncertainty 9
Figure 2. PRISMA graphic for literature search for key questions 1, 2 and 3 26
Figure 3. Evidence Continuum in Implementation Science 106

Appendixes

Appendix A. Search Strategy
Appendix B. Excluded Studies
Appendix C. Risk of Bias Tables
Appendix D. Evidence Tables for Key Question 1
Appendix E. Evidence Tables for Key Question 2
Appendix F. Evidence Tables for Key Question 3

Executive Summary

Introduction

The Agency for Healthcare Research and Quality (AHRQ) sponsors research to improve the quality, effectiveness, and safety of health care in the United States. Evidence reports and technology assessments generated through AHRQ's Effective Health Care (EHC) Program provide science-based information about common, relevant health conditions and technologies to serve the needs of patients, clinicians, insurance payers, and other end users. Findings from clinical, health services, and comparative effectiveness studies—especially as assembled for systematic reviews and similar documents—need to be communicated and disseminated effectively to influence optimal and timely practice and health policies.[1]

Because systematic reviews evaluate multiple studies, they are inherently complex. Nuanced descriptions of benefits, harms, strengths of evidence, and uncertainties make evidence reports difficult for many people to understand. Evidence reports typically target scientific researchers in related fields, rather than patients, clinicians, or policy makers, who are ultimately responsible for making health related decisions and setting policy. Ensuring that research evidence reaches clinicians and patients in easy-to-understand formats is critical to the success of evidence-based research and practice.

Given AHRQ's mission, evaluating the effectiveness of strategies to make evidence report findings widely accessible and understood is critical. Findings from this review to evaluate the comparative effectiveness of communication techniques and dissemination strategies will inform efforts to make research easily accessible for patients, clinicians, and policy makers.

Objective

This systematic review has three related components; all focus on promoting informed decisions about health related behaviors and decisions among patients and clinicians. First, the review addresses the comparative effectiveness of **communicating evidence** in various contents and formats that increase the likelihood that target audiences will both understand and use the information. Second, it examines the comparative effectiveness of a variety of approaches for **disseminating evidence** that increase the likelihood of evidence adoption and use. Third, it examines the comparative effectiveness of ways to **communicate uncertain** health-related evidence that increase the likelihood that members of different target audiences will understand and use the information. Due to the complexities of our questions, we present our work as three separate systematic reviews—one for communication, one for dissemination, and a third for uncertainty—each addressing a separate key question. Combined, these three separate reviews inform how to best communicate and disseminate evidence.

Communication Strategies to Promote the Use of Health Care Evidence

Government agencies and institutions, advocacy groups, media organizations, researchers, and other interested stakeholders can all carry out communication activities. They use various strategies to communicate evidence so that target audiences can understand it better and increase the probability that recipients pay attention to the messages conveyed.[2] Health communication, defined as the study and use of communication strategies to inform and influence individual and

community decisions that affect health,[3] is increasingly recognized as a necessary element of efforts to improve personal and public health.

. To our knowledge, no overarching framework exists to guide the part of our review focused on communication strategies. Multiple systematic reviews, however, have explicated key communication strategies that are of interest to the field. They include four core constructs:

1. **Tailoring the message—**Communication designed for an individual based on information from the individual.
2. **Targeting the message to audience segments—**Communication designed for subgroups based on group membership or characteristics such as age, sex, race, cultural background, language, and other "psychographic" characteristics such as a person's attitudes about a particular subject matter.
3. **Using narratives—**Communication delivered in the form of a story, testimonial, or entertainment education.
4. **Framing the message—**Communication that conveys the same messages in alternate ways (e.g., what is gained or lost by taking an action or making a choice).

The four strategies are distinct but could be used simultaneously.

Key Question 1

a. What is the comparative effectiveness of communication strategies to promote the use of health and health care evidence by patients and clinicians?
b. How does the comparative effectiveness of communication strategies vary by patients and clinicians?

Dissemination Strategies to Promote the Use of Health Care Evidence

Dissemination is the active and targeted distribution of information or interventions via determined channels using planned strategies to a specific public health or clinical practice audience.[4-6] Dissemination has been characterized as a necessary but not sufficient antecedent of adoption and implementation. In contrast to diffusion, which is a passive, informal process, dissemination is a formal, planned process with the intent of spreading knowledge and associated evidence-based interventions to stimulate adoption and enhance the integration of the evidence, information, or intervention (or combinations of these) into routine practice.[4,5,7-9]

Evidence dissemination has several very broad goals relating to evidence and information:

1. **Increase reach to a variety of audiences—**Distributing evidence widely to many audiences and across many settings increases the reach of information (e.g., postal and electronic mail; electronic/digital, social, and mass media)
2. **Increase motivation to use and apply such information—**Increasing interest in the evidence through champions (aka a cheerleader), opinion/thought leaders, or social networks
3. **Increase ability to use and apply evidence**—Providing additional resources about the evidence, such as how it can be incorporated into current practice or specific suggestions for change, to enhance a traditional dissemination strategy (e.g., by also providing additional resources or information; skills-building efforts)

In addition, it is also common practice to combine multiple dissemination strategies (labeled as *multi-component* strategies in this review) to address a combination of reach, ability or motivation goals.

Key Question 2

a. What is the comparative effectiveness of dissemination strategies to promote the use of health and health care evidence for patients and clinicians?
b. How does the comparative effectiveness of dissemination strategies vary by patients and clinicians?

Explaining Uncertain Evidence

Uncertainty is inherent in health care and evidence about health care.[10] It stems from multiple sources, including imperfect knowledge about scientific evidence, patients' preferences and circumstances, and how to apply judgment in decisionmaking.[10-13] Uncertainty interferes with both patients' and physicians' ability to derive appropriate meaning about illness, diagnostic tests, treatments, and prognosis and to use this information in meaningful ways. In the context of evidence translation, uncertainty creates multiple challenges including: (1) determining whether preventive services and treatments should be implemented in clinical practice, (2) determining for whom and in what settings preventive services and treatments should be implemented, and (3) communicating evidence so that end-users can make informed decisions.

For our review, we synthesized the current literature to present our own framework of uncertainty as it relates to evidence translation.[10-16] It encompasses the concepts of uncertainty relative to the current scheme for grading the strength of evidence for AHRQ's Evidence-based Practice Center (EPC) program, including risk of bias, consistency, directness, and precision.[17] It also includes the broader ideas of net benefit (i.e. whether is more benefit than harm or vice versa) and applicability.[18,19]

1. **Overall strength of evidence**—Degree of confidence that the estimates of effects are correct and represent the true effect.
2. **Risk of Bias—**Degree to which individual studies are protected from systematic errors or bias.
3. **Consistency—**Degree to which studies present similar findings—in either direction of effect or magnitude of effect (or both).
4. **Precision—**Degree of random error surrounding an effect estimate with respect to a given outcome; such studies express dispersion around a point estimate of risk, such as a confidence interval.
5. **Directness—**Degree to which the evidence either directly links the interventions to the outcome of interest or directly makes the comparison of interest.
6. **Net benefit—**Balance or trade-offs in benefits and harms for prevention or treatment services.
7. **Applicability—**Reflects whether an intervention is expected to have the same effect in populations and settings where it was not studied, but might be applied.

Key Question 3

a. What is the comparative effectiveness of different ways of explaining uncertain health and health care evidence to patients and clinicians?

Analytic Framework

We present our analytic framework in Figure A. As noted in the box to the far left, we examined studies that used research-based evidence as the source of information for their communication strategies (KQ 1) and dissemination strategies (KQ 2). For KQ1 and 2, we specifically defined research-based evidence as evidence that has been assembled, reviewed, and presented by evidence developers and that has been used to make recommendations. For KQ3, however, we accepted any type of evidence presented given the paucity of overall literature.

Figure A. Analytic framework for communicating and disseminating evidence and explaining uncertainty

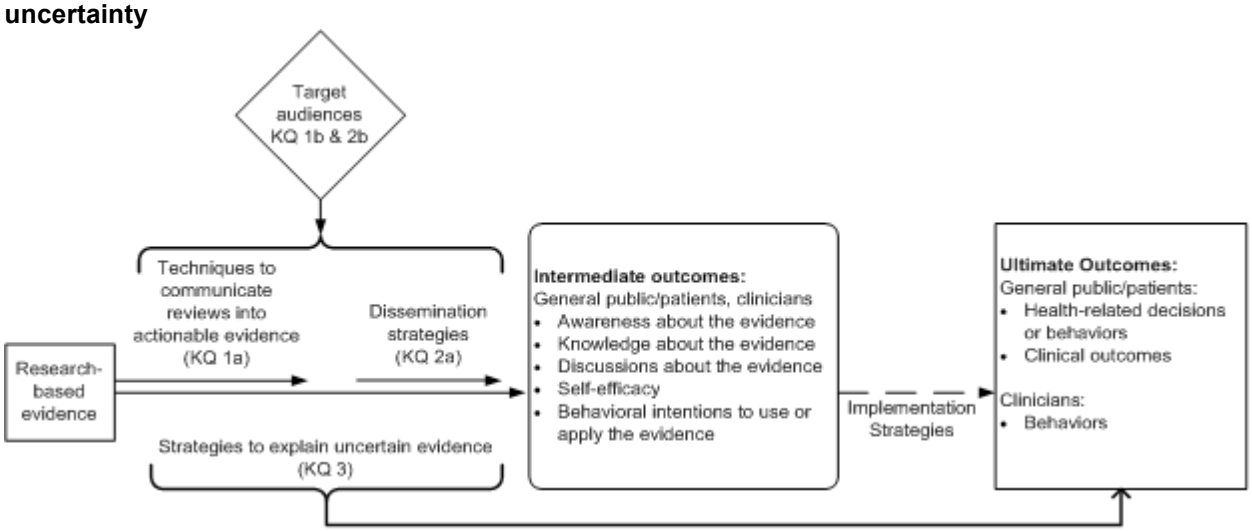


Figure A also details outcomes that we included in our review. We included studies that examined both intermediate and ultimate (distal) outcomes. Intermediate outcomes can be awareness of the evidence, knowledge of the evidence, discussions about the evidence, self-efficacy (or one's confidence) to use the evidence, and behavioral intentions to use or apply the evidence. Ultimate outcomes include be the following: for patients—health-related decisions or behaviors and clinical outcomes; for clinicians—behaviors. We expected that most studies focused on intermediate outcomes because they occur sooner and, thus, are more practical to study.

Our inclusion and exclusion criteria address both the PICOTS model (population, interventions, comparators, outcomes, timeframes, and settings; see Introduction chapter) and other important study design and publication issues. Table 6 of the full report shows the inclusion/exclusion criteria common to all three KQs. There was also specific inclusion criteria applied to admissible research evidence for KQ1 and KQ2 (shown in Table 7 of full report) and other KQ-specific inclusion/exclusion criteria (shown in Tables 8, 9, and 10, respectively of full report).

Methods

Literature Search and Retrieval Process

We systematically searched, reviewed, and synthesized the scientific evidence for each KQ separately. Databases included MEDLINE®, the Cochrane Library, Cochrane Central Trials Registry, PsychInfo, and the Web of Science. We did not conduct additional searches for gray literature.

We used a variety of medical subject headings (MeSH terms) and major headings, free-text and title and abstract text-word searches. Search results were limited to studies on humans published from 01/01/2000 to May 15, 2012 for communication and dissemination given the time frames covered by prior systematic reviews, and from 01/01/1966 onward to May 15, 2012 for uncertainty given the lack of prior reviews on this specific topic.

We hand-searched bibliographies of included articles. In addition, in an effort to avoid retrieval bias, we manually searched the reference lists of landmark studies and background articles on this topic to look for any relevant citations that electronic searches might have missed.

Article Review and Data Abstraction

We used standard EPC methods for dual review of abstracts and full text of articles to determine article inclusion. After determining article inclusion, one reviewer entered data about studies into evidence tables and a second, senior member of the team reviewed all abstractions against the accompanying article(s) for completeness and accuracy.

Risk of Bias Assessment of Individual Studies

Two reviewers independently rated the risk of bias of studies (low, medium, or high) using criteria designed to detect selection bias (including attrition bias), measurement bias (such as performance bias, detection bias), confounding, and inadequate power. We also assessed potential biases in reporting. Reviewers resolved all disagreements about quality ratings by discussion and consensus or by consulting a third, senior member of the team. We did not retain studies of high risk of bias for analysis, presentation in the results chapters or strength of evidence grading. Studies with a high risk of bias were those with at least one major flaw that was likely to cause significant bias and thus might have invalidated the results. Major flaws preclude the ability to draw causal inferences between the intervention and the outcome.

Data Synthesis and Grading Strength of Evidence

Studies included in our review compared a wide range of interventions and plethora of outcomes; they were sufficiently heterogeneous to preclude meta-analysis. Thus, we synthesized the data qualitatively by KQ. We paid particular attention to moderators of study effects as a way to explain any seemingly disparate findings. Possible moderators of interest for all key questions included risk of bias, study size, and target audience.

The investigative team jointly discussed and graded the overall body of literature and generated recommendations for future research. We graded the strength of evidence on the basis of guidance established for the EPC program.[20,21] The EPC approach incorporates four required domains: risk of bias (including study design and aggregate quality), consistency, directness, and precision of the evidence. Two reviewers independently rated the four domains for each

intervention for each key outcome (listed in the analytic framework depicted in Figure A); conflicts were resolved by group consensus. The research group derived the overall strength of evidence grade (resolving conflicts in the same way).

Results

Search Results and Included Studies

We identified 3,917 articles from all sources (after removing duplicates) for all three key questions. After applying the inclusion and exclusion criteria, 374 articles were retained for full text review. The majority of the full-text articles were classified to one or more Key Question (KQ): 93 articles pertained to KQ1, 153 articles pertained to KQ2 (one of which overlapped with KQ1), 80 articles pertained to KQ3, and 49 articles were classified as overlap. Each overlap article potentially applied to two or more KQs and was not classified into one KQ category. Of the full-text articles, we excluded 321, leaving 53 articles for data abstraction. Six articles (representing five studies) are relevant to KQ 1; 40 articles (representing 37 studies) are relevant to KQ 2; and eight articles (representing seven studies) are relevant to KQ 3.

Key Question 1: Communication Strategies

Of the 93 articles pertinent to KQ1, we retained six articles (representing five trials) after full-text review that met inclusion criteria and report on five trials.[22-27] Two trials addressed tailoring versus targeting; one addressed a combination of targeting and tailoring versus tailoring only, one addressed framing versus targeting; and one addressed framing versus narratives. Three studies were RCTs, one was a factorial RCT, and one was a randomized trial. The investigators tested these interventions in patient populations in the United States; no studies included clinicians. Sample sizes ranged from 174 participants to 5,500 participants. Outcomes studied included self-reported breast and cervical cancer screening and dietary behaviors.

Key Points

- **Targeting versus tailoring:** Findings were either counterintuitive or non-significant for the two studies that compared targeting with tailoring. In both studies, investigators hypothesized that the tailored version of the intervention would have a greater effect on screening rates than the targeted version. However, the *targeted* version was associated with greater likelihood of self-reported screening relative to the *tailored* version in one study. The authors attributed this unexpected finding to a possible "boomerang" (i.e., reverse) effect or possibly a lack of meaningful difference between the two versions of the intervention. In another trial, there were no significant differences in dietary outcomes between those receiving the targeted or tailored version of the intervention. Adding a lay health worker to the tailored version of the intervention did not improve dietary outcomes (two trials; low SOE).
- **Targeting and tailoring versus targeting only**. Investigators found no statistically significant differences when they targeted an intervention to the subpopulation *and* personally tailored it to each study participant compared to a version of the intervention that was only targeted. They attributed the lack of differential impact to a possible 'ceiling effect' in the study population given the fairly high baseline screening rates (of about 80 percent) (one trial; low SOE).

- **Use of multiple communication strategies simultaneously:** In several cases, investigators used some combination of the four communication strategies when developing their interventions instead of comparing only a single strategy with another single strategy. Because comparisons were not one-to-one, it was more challenging to isolate the effects of each strategy.
- **Combining communication strategies with channel variation.** In one trial, investigators enhanced the communication strategy by *also* varying the communication channel for the intervention (i.e., using a lay health worker). While this tactic creates the potential for a more powerful effect, it is also complicates determining the effect of each strategy relative to the other.
- **Variation in use of strategies for patients versus clinicians.** None of the trials that met our review criteria use any of the four communication strategies with clinicians, therefore, we were unable to address KQ1b which addressed use of the strategies with clinicians.

Key Question 2: Dissemination Strategies

We included 37 studies that focused on evidence dissemination to clinicians or patients (broadly defined) and that used strategies focusing on increasing the reach, ability, motivation or used a multi-component approach to enhance health-related decisions or behaviors, clinical outcomes or knowledge. We divided the trials by dissemination strategies and by outcomes for clinicians and patients. Some trials compared strategies directly with each other (e.g., ability strategies vs. motivation strategies) and can be regarded as head-to-head trials for comparative effectiveness analyses. Some trials compared two or more strategies to a usual care or no-treatment control group, but we included them in our analysis if we believed that we might glean information from indirect statistical comparisons about the relative effectiveness of one strategy versus another. We note cases in the summary tables where there were not significant tests or confidence intervals for the direct comparison.

The 37 trials reported a wide variety of primary and secondary outcomes that spanned a range of health-related or clinical problems. Nineteen trials addressed disseminating information to clinicians, 12 to patients, and 6 to patients versus clinicians. Of trials addressing disseminating evidence to clinicians, 7 were RCTs and 12 were cRCTs. One trial compared reach to motivation strategies, 6 reach to ability strategies, 3 motivation to ability, and 14 reach, motivation, or ability to multicomponent strategies. Of 12 trials addressing disseminating evidence to patients, 11 were RCTs and one was a fRCT. Six trials compared reach to motivation strategies, 6 various reach strategies, and 7 reach or motivation to multicomponent strategies. Of trials addressing disseminating evidence to patients versus physicians, 2 were RCTs, 4 were cRCTs and one was fRCT. All seven studies compared reach motivation, or ability to multicomponent strategies. The trials were conducted in the United States, Canada, England, Germany, Finland, the Netherlands, Scotland, and Spain. Sample sizes ranged from 114 participants to 3,293 participants. For the cRCTs, cluster sizes ranged from 9 to 249.

Key Points: Disseminating Evidence to Clinicians

- Ability strategies were not more effective than reach strategies related to clinician behavior (4 trials, low SOE due to non-significant findings across all studies).

- One study found that motivation strategies may be more effective than reach strategies for influencing clinicians' behaviors (1 trial; insufficient evidence).
- Multi-component strategies that address a combination of reach, ability, or motivation appear to be more effective than one strategy alone for affecting clinicians' behaviors, particularly guideline adherence (6 trials; moderate SOE) and for clinical outcomes although many comparisons examining clinical outcomes were not significant (4 trials, low SOE).
- The strength of evidence is insufficient for most comparisons related to clinical outcomes and knowledge for clinicians because we had only single trials in each case.

Key Points: Disseminating Evidence to Patients

- Evidence was inconsistent for determining the benefit of reach, ability, motivation, or multi-component approaches for patients focused on changing health-related decisions and behaviors. (12 trials; insufficient SOE).
- Evidence was insufficient for determining the benefit of reach, ability, motivation or multi-component approaches for patients focused on changing clinical outcomes. (2 trials; insufficient SOE due to one trial in each category).
- Evidence was insufficient for determining the benefit of reach, ability, motivation or multi-component approaches for patients focused on changing knowledge outcomes (3 trials; insufficient SOE due to inconsistent findings or one trial in a category).

Key Points: Disseminating Evidence to Patients and Clinicians

- Evidence was inconsistent for determining the benefit of reach, ability, motivation, or multi-component strategies that target both providers and patients for health-related decisions and behaviors (4 trials; insufficient SOE).
- Evidence was inconsistent for determining the benefit of reach, ability, motivation, or multi-component strategies that target both providers and patients for health-related decisions and behaviors or clinical outcomes (1 trial in each category; insufficient SOE).

Key Question 3: Uncertainty

We found seven unique studies that met our inclusion criteria, had low or moderate risk of bias, and examined alternate ways to communicate the overall strength of evidence, and the precision, directness, and net benefit of evidence. We found no eligible studies on risk of bias, consistency, or applicability. Of these studies, one was a RCT, four were factorial RCTs, one a non-controlled trial, and one a quasi-experimental study. One reported on the effects of alternate wordings of overall strength of evidence.[28] Three studies reported on various presentations of precision;[29,30] one tested alternative ways of communicating directness;[31] and three investigated different ways of communicating net benefit (with some studies making more than one comparison).[31-35] Three studies reported the effects of alternate non-numeric presentations of uncertainty;[28,31,33] two on alternate numeric presentations;[10,30] one on numeric versus graphical presentations;[10] one on alternate graphical presentations;[10] and 1 on framing.[34]

Interventions were tested in study populations in the United States, Canada, and Switzerland. Sample sizes ranged from 135 participants to 2,234 participants. Outcomes studied included knowledge, perceived risk, accuracy of perceived risk, appropriate choices regarding care (e.g., selecting medications; obtaining screening), and decision satisfaction.

Key Points

- **Communicating strength of evidence:** One study found that no single non-numeric recommendation scheme (e.g. GRADE, NICE, ACCP) was superior to others for all types of recommendations (strong for, weak for, weak against, strong against). Different ways of wording recommendations intended to convey strong recommendations resulted in only moderate guideline-concordant care with no wording superior to others. Weak recommendation wording resulted in higher guideline concordant care with "we suggest" and "we conditionally recommend against" wording performing better than other wording for and against care. (one study; insufficient SOE).
- **Communicating precision:** One study found that presenting numeric risks as point estimates versus 95% confidence intervals (CIs) or changing the format of 95% CIs (numeric versus graphical) did not significantly influence perceived risk of colon cancer (one study; insufficient SOE) or the accuracy of perceived risk of breast cancer recurrence (one study, insufficient evidence) Using clean versus blurry bar graphs to convey information about uncertainty did not produce different levels of perceived risk (one study; low SOE)
- **Communicating directness:** Choice of a cholesterol medication with direct evidence of net benefit was better for patients receiving non-numeric statements that provided either directive advice (i.e. "ask for the drug that…") or factual information about the evidence than for patients receiving usual care. However, medication choices did not differ by type of instruction (one study; low SOE).
- **Communicating net benefit:** Choice of a heartburn medication that was more likely to have net benefit because of its longer history of evidence about safety was better for consumers receiving non-numeric statements that provided either directive advice or factual information rather than usual care, but medication choices did not differ by type of instruction (one study; low SOE). Receiving additional non-numeric information about benefits had little effect on refusals of cancer screening tests, but receiving more non-numeric information on harms significantly increased test refusals and significantly decreased decision satisfaction (all low SOE). Compared with usual care, giving men prostate cancer screening information alone or framed in the context of information about other, more beneficial screening services significantly increased prostate cancer knowledge and the proportion of individuals reporting active involvement in decision-making. However, giving prostate cancer screening information alone increased screening, whereas giving it framed in the broader context of more beneficial services decreased screening (all low SOE).

Discussion

This report presented three separate, but topically related, systematic reviews. The overarching topic involves providing health-related evidence effectively to patients and clinicians. Specifically, we were asked to examine various strategies for communicating and disseminating evidence to these target audiences. Finally, we were charged with exploring ways to explain uncertainty in evidence.

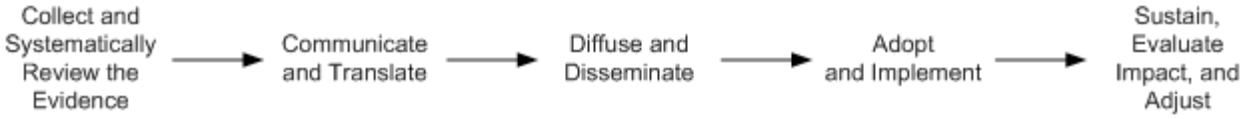
Many aspects of this review cut across more than one KQ, and some across all three KQs. Below we set findings from our research into the broader context of evidence translation and highlight key cross-cutting issues that, if addressed, might advance the field. We also discuss limitations within each question and of our overall review process that should be considered in

interpreting our results. Finally, we highlight the implications of our work for future research and ramifications for patients, clinicians, and other stakeholders and end users.

Evidence in the Broader Context

- **The Evidence Continuum.** In the context of our review, we view the evidence as moving along a continuum beginning with its collection and systematic review followed by communicating and translating it for audiences as needed (see Figure B). The communication and translation process is often comingled with the diffusion (passive spread) and dissemination (active spread) of the information. Our review included only the second and third phases in the evidence continuum. Some trials seemed to conflate communication and dissemination, perhaps not surprisingly given how difficult cleanly defining these concepts can be. Several other trials also seemed to mix or merge dissemination with implementation. This conceptual overlap complicated our analysis in at least two stages: creating meaningful classifications of strategies reported in the literature and examining appropriate relevant outcomes for those strategies.

Figure B. Evidence continuum in implementation science



Cross-Cutting Issues to Address in Advancing the Field

- **Definitions of Concepts and Terms.** Consensus is lacking regarding definitions of key terms pertinent to this review and the research efforts more generally. We saw this lack of consensus across studies especially for definitions of three key terms: dissemination, adoption, and implementation. Greater unity in the field in terms of concepts and terms would be beneficial. With respect to KQ2, the lack of consistency in how the field references and classifies dissemination strategies hampered our efforts to classify a strategy into one of our domain groupings.
- **Use of Theoretical Frameworks and Models.** Many studies (but not all) lacked any apparent theoretical or conceptual framework to inform or organize the research questions and focus interventions on essential processes of behavioral and systems change.
- **Methodological considerations.** In the included trials, investigators sometimes did not fully capitalize on the available data, which may partly explain why many of our included studies showed little or no effect of specific intervention strategies. Many of the studies only employed descriptive statistics and did not capitalize on more recent methodological advances (e.g., multi-level modeling) that could have improved their analytic approach. Other studies did not factor in potentially important moderating variables such as self-efficacy and health literacy.

Limitations of the Literature Specific to Key Question

Examples of limitations for Key Question 1 trials included:

- Evidence base for addressing comparisons of communication strategies of interest was extremely sparse (i.e., only five trials of direct comparisons);
- Trials almost exclusively focused on breast and cervical cancer screening interventions. The evidence basis for breast cancer screenings has changed in the recent past. As new evidence emerges in the media the result can be confusion among patients and may produce potential interference with the impact of interventions;
- All trials used convenience samples, so unmeasured confounding may exist because of selection bias with the sample;
- All trials used self-reported data, which can be subject to social desirability bias.

Examples of limitations for Key Question 2 trials included:

- Trials often confounded the mode of distribution with other variables. Therefore, we could not tease apart the effect of mode, channel, and other variables on the outcome of interest;
- Studies inconsistently reported and/or described all aspects of study design and conduct, especially for interventions.
- Many studies did not consistently compare strategies directly to each other, but instead compared to a usual care or control condition or at times made direct comparisons for only some outcomes. This limited our ability to draw conclusions about the comparative effectiveness of one approach versus another; and
- The included studies were very heterogeneous with regard to the behaviors, outcomes, targeted populations, and dissemination strategies used. The resulting heterogeneity reflects a commonly encountered attribute of dissemination research. To address this heterogeneous and complicated body of work, we classified the trials in broad terms. Nonetheless, this effort still left too few studies in some categories with which to make meaningful conclusions about the relative impact of a particular dissemination strategy.

Examples of limitations for Key Question 3 trials included:

- Trials did not directly test alternative ways to communicate the uncertainty concepts that are relevant to evidence about health and health care. Few studies addressed any type of uncertainty of interest and none examined ways to communicate risk of bias, consistency across studies, or applicability
- When acceptable studies were present, we determined that they manipulated relatively limited comparisons. For instance, few alternate wordings were tested for communicating strength of evidence, and few graphical presentations were tested for communicating precision.
- Samples sizes often small and non-representative

Future Research

Research teams should try to address both the conceptual and the study limitations noted both in sections addressing each key question and here above. This includes:

- Relying more on accepted theoretical constructs and models when designing interventions and studies;

- Conducting some prior needs assessments with target audiences, focusing on audience subgroups with greatest needs;
- Using more sophisticated and well-powered study designs;
- Describing and defending choices of intermediate and ultimate outcomes;
- Applying modeling or other advanced statistical and analytic techniques to account for confounders, interactions, and similar complications in data, and addressing temporal aspects of outcomes; and
- Thoroughly describing all aspects of study design and conduct, especially for interventions.

Implications of This Report for Clinicians and Policymakers

Clinicians and policy makers are in unique positions to use effective communication and dissemination strategies to promote and accelerate the adoption of the evidence base for improving health and healthcare. In many studies, clinicians and policymakers are also members of the intended audience for communication and dissemination strategies designed to alert them to evidence necessitating change in clinical practice and policies guiding it. This review sought to identify effective strategies for communicating and disseminating evidence so that patients and caregivers, practicing clinicians, relevant professional and consumer organizations, purchasers of health care, and others can understand and choose whether to use available evidence as a guide for their clinical practice and policy choices.

Our findings offer some guidance for clinicians and policymakers as to the most effective strategies for communicating and disseminating evidence, but leave many questions unanswered. While clinicians and policymakers may use our findings to guide choice of a specific communication and/or dissemination strategy, they should also carefully consider other factors shown to affect awareness, adoption and use of evidence in various settings and by individuals working in or receiving services in those settings.

Patient-centered communication is essential to the provision of patient-centered care and especially so in the face of uncertain evidence about treatment options.[36] We examined the impact of strategies to communicate uncertain evidence including approaches to communicating strength of evidence, precision and directness of risk estimates, and net benefit. Our findings show no consistent impact of these strategies on changing patient outcomes such as knowledge, perceived risk, accuracy of perceived risk, appropriate choices regarding care (e.g., selecting medications; obtaining screening), and decision satisfaction. As clinicians decide how to communicate evidence to their patients, they should consider multi-faceted, multi-directional sharing of information between themselves and their patients that takes into account health literacy, cultural and other factors affecting an individual patient's ability to process and retain information, and use it to make an informed decision. This interaction also hinges on individual clinician knowledge of evidence, organizational policies and standard protocols, and resources available to both the clinician and patient.

The lack of research evidence to inform communication and dissemination of evidence, including uncertain evidence impedes timely clinician, patient and policymaker awareness, uptake and use of evidence to improve the quality of care. Expanding investment in the relatively new field of communication and dissemination research is critical to the identification of strategies to accelerate the translation of comparative effectiveness research into community and clinical practice and the direct benefit of patient care.

More research is needed to better understand the current barriers to translating the findings of comparative effectiveness research into community and clinical practice.[37] Ongoing funding for interdisciplinary communication and dissemination sciences research is needed to promote the uptake and use of evidence and ensure quality of care.

Conclusions

In closing, this was the first systematic review that attempted to compare the effectiveness of communication strategies and to look at communicating uncertainty. Finding the appropriate "comparative" studies was challenging. The number of eligible studies was more limited for KQ1 and KQ3, but more substantial for KQ2. The review provides a helpful foundation in setting the research agenda to address key gaps in the literature.

References

1. Helfand M, Tunis S, Whitlock EP, et al. A CTSA agenda to advance methods for comparative effectiveness research. Clin Transl Sci. 2011 Jun;4(3):188-98. PMID: 21707950.
2. McGuire WJ. Personality and attitude change: An information processing theory. In: Greenwald AG, Brock TC, Ostrom TM, eds. Psychological foundations of attitudes. San Diego, CA: Academic Press; 1968:171-96.
3. U.S. Department of Health and Human Services. Health.gov. Washington, DC: Office of the Assistant Secretary for Health, Office of the Secretary, U.S. Department of Health and Human Services. http://www.health.gov/communication/resources/Default.asp. Accessed April 5, 2012.
4. Rabin BA, Brownson RC, Haire-Joshu D, et al. A glossary for dissemination and implementation research in health. J Public Health Manag Pract. 2008 Mar-Apr;14(2):117-23. PMID: 18287916.
5. Brownson RC, Colditz GA, Proctor EK. Dissemination and implementation research in health. New York, NY: Oxford University Press; 2012.
6. Tabak RG, Khoong EC, Chambers DA, et al. Bridging research and practice: models for dissemination and implementation research. Am J Prev Med. 2012 Sep;43(3):337-50. PMID: 22898128.
7. National Institutes of Health. NIH Conference. Building the Science of Dissemination and Implementation in the Service of Public Health. 2007 Sep 10-11. http://obssr.od.nih.gov/di2007/about.html. Accessed February 21, 2012.
8. Lomas J. Diffusion, dissemination, and implementation: who should do what? Ann N Y Acad Sci. 1993 Dec 31;703:226-35; discussion 35-37. PMID: 8192299.
9. Greenhalgh T, Robert G, Macfarlane F, et al. Diffusion of innovations in service organizations: Systematic review and recommendations. Milbank Q. 2004;82(4):581-629. PMID: 15595944.
10. Han PK, Klein WM, Arora NK. Varieties of uncertainty in health care: a conceptual taxonomy. Med Decis Making. 2011 Nov-Dec;31(6):828-38. PMID: 22067431.
11. Babrow AS, Kasch CR, Ford LA. The many meanings of uncertainty in illness: toward a systematic accounting. Health Commun. 1998;10(1):1-23. PMID: 16370987.
12. Djulbegovic B. Articulating and responding to uncertainties in clinical research. J Med Philos. 2007 Mar-Apr;32(2):79-98. PMID: 17454416.
13. Rowe WD. Understanding Uncertainty. Risk Anal. 1994;14(5):743-50.
14. Politi MC, Han PK, Col NF. Communicating the uncertainty of harms and benefits of medical interventions. Med Decis Making. 2007 Sep-Oct;27(5):681-95. PMID: 17873256.
15. Mishel MH. Uncertainty in illness. Image--the journal of nursing scholarship. 1988 Winter;20(4):225-32. PMID: 3203947.
16. Tannert C, Elvers HD, Jandrig B. The ethics of uncertainty. In the light of possible dangers, research becomes a moral duty. EMBO reports. 2007 Oct;8(10):892-6. PMID: 17906667.
17. Owens DK, Lohr KN, Atkins D, et al. Grading the Strength of a Body of Evidence When Comparing Medical Interventions. Methods Guide for Effectiveness and Comparative Effectiveness Reviews. Rockville, MD: Agency for Healthcare Research and Quality; 2008.
18. Harris RP, Helfand M, Woolf SH, et al. Current methods of the US Preventive Services Task Force: a review of the process. Am J Prev Med. 2001 Apr;20(3 Suppl):21-35. PMID: 11306229.
19. Sawaya GF, Guirguis-Blake J, LeFevre M, et al. Update on the methods of the U.S. Preventive Services Task Force: estimating certainty and magnitude of net benefit. Ann Intern Med. 2007 Dec 18;147(12):871-5. PMID: 18087058.

20. Agency for Healthcare Research and Quality. Methods Guide for Effectiveness and Comparative Effectiveness Reviews. Agency for Healthcare Research and Quality AHRQ Publication No. 10(11)-EHC063-EF. Rockville, MD: February 2011. Chapters available at: www.effectivehealthcare.ahrq.gov.

21. Owens DK, Lohr KN, Atkins D, et al. AHRQ series paper 5: Grading the strength of a body of evidence when comparing medical interventions--Agency for Healthcare Research and Quality and the Effective Health-Care program. J Clin Epidemiol. 2010 May;63(5):513-23. PMID: 19595577.

22. Cox D, Cox AD. Communicating the consequences of early detection: The role of evidence and framing. J Market. 2001;65(3):91-103.

23. Schneider TR, Salovey P, Apanovitch AM, et al. The effects of message framing and ethnic targeting on mammography use among low-income women. Health Psychol. 2001;20(4):256-66. PMID: 11515737.

24. Jibaja-Weiss ML, Volk RJ, Kingery P, et al. Tailored messages for breast and cervical cancer screening of low-income and minority women using medical records data. Patient Educ Couns. 2003 Jun;50(2):123-32. PMID: 12781927.

25. Vernon SW, del Junco DJ, Tiro JA, et al. Promoting regular mammography screening II. Results from a randomized controlled trial in US women veterans. J Natl Cancer Inst. 2008 Mar 5;100(5):347-58. PMID: 18314474.

26. del Junco DJ, Vernon SW, Coan SP, et al. Promoting regular mammography screening I. A systematic assessment of validity in a randomized trial. J Natl Cancer Inst. 2008 Mar 5;100(5):333-46. PMID: 18314473.

27. Elder JP, Ayala GX, Campbell NR, et al. Long-term effects of a communication intervention for Spanish-dominant Latinas. Am J Prev Med. 2006 Aug;31(2):159-66. PMID: 16829333.

28. Akl EA, Guyatt GH, Irani J, et al. "Might" or "suggest"? No wording approach was clearly superior in conveying the strength of recommendation. J Clin Epidemiol. 2012 Mar;65(3):268-75. PMID: 22075112.

29. Kylat Ranjit I, Ohlsson A. Recombinant human activated protein C for severe sepsis in neonates. Cochrane Database of Systematic Reviews: John Wiley & Sons, Ltd; 2012.

30. Brewer NT, Richman AR, DeFrank JT, et al. Improving communication of breast cancer recurrence risk. Breast Cancer Res Treat. 2012 Jun;133(2):553-61. PMID: 21964579.

31. Schwartz LM, Woloshin S. Communicating uncertainties about prescription drugs to the public: a national randomized trial. Arch Intern Med. 2011 Sep 12;171(16):1463-8. PMID: 21911629.

32. Perneger TV, Schiesari L, Cullati S, et al. Does information about risks and benefits improve the decision-making process in cancer screening - randomized study. Cancer Epidemiol. 2011 Dec;35(6):574-9. PMID: 21622043.

33. Perneger TV, Cullati S, Schiesari L, et al. Impact of information about risks and benefits of cancer screening on intended participation. Eur J Cancer. 2010 Aug;46(12):2267-74. PMID: 20466537.

34. McCormack L, Treiman K, Bann C, et al. Translating medical evidence to promote informed health care decisions. Health Serv Res. 2011 Aug;46(4):1200-23. PMID: 21352225.

35. McCormack LA, Bann CM, Williams-Piehota P, et al. Communication message strategies for increasing knowledge about prostate cancer screening. J Cancer Educ. 2009;24(3):238-43. PMID: 19526414.

36. McCormack LA, Treiman K, Rupert D, et al. Measuring patient-centered communication in cancer care: A literature review and the development of a systematic approach. Soc Sci Med. 2011 Apr;72(7):1085-95. PMID: 21376443.

37. Avorn J, Fischer M. 'Bench to behavior': translating comparative effectiveness research into improved clinical practice. Health Aff (Millwood). 2010 Oct;29(10):1891-900. PMID: 20921491.

Introduction

Background

Rationale and Relevance for Conducting the Systematic Review

The Agency for Healthcare Research and Quality (AHRQ) sponsors research to improve the quality, effectiveness, and safety of health care in the United States. Evidence reports and technology assessments generated through AHRQ's Effective Health Care (EHC) Program provide science-based information about common, relevant health conditions and technologies to serve the needs of patients, clinicians, insurance payers, and other end users. Findings from clinical, health services, and comparative effectiveness studies—especially as assembled for systematic reviews and similar documents—need to be communicated and disseminated effectively to influence optimal and timely practice and health policies.[1]

Because systematic reviews evaluate multiple studies, they are inherently complex. Nuanced descriptions of benefits, harms, strengths of evidence, and uncertainties make evidence reports difficult to understand for many people. Evidence reports are typically targeted at scientific researchers in related fields, rather than at the patients or clinicians who ultimately make health related decisions. Ensuring that research evidence is delivered to these audiences in easy-to-understand formats is critical to the success of evidence-based research. Common communication and dissemination barriers, including not seeing or being exposed to the information, can impede its use in decisionmaking.[2-4]

Given AHRQ's mission, a critical goal is to evaluate the effectiveness of strategies to make evidence report findings widely available and techniques to ensure that such findings are correctly understood. By evaluating the comparative effectiveness of communication techniques and dissemination strategies, this review will inform efforts to make research easily accessible for patients and clinicians.

Due to the complexities of our questions, we present our work as three separate systematic reviews – one for communication, one for dissemination, and a third for uncertainty—each addressing a separate key question. Combined, these three separate reviews inform how to best translate and disseminate evidence.

Terminology and Definitions

Transforming scientific evidence for its use in practice, commonly known as *research translation*, involves many processes and strategies. Investigators must conduct high-quality studies; experts must synthesize and summarize these bodies of evidence, often in the form of systematic reviews of comparative effectiveness. Authors of material typically presented in complex and technical jargon must alter it to simpler language and actionable steps that potential end users will find easier to understand; authors or organizations must disseminate such documents to those audiences; and, finally, providers and others must incorporate the information into existing health care processes and systems to improve health.

The terminology for each of these steps overlaps considerably. Our review focuses on the communication and dissemination of evidence about health and health care and examines effective ways to present associated uncertainty. Table 1 lists four key definitions to help readers understand the scope of our review and the concepts that we will use throughout it.

Table 1. Definitions of concepts relevant for this review

Concept or Construct	Definition As It Relates to Health and Health Care
Health communication	The study and use of communication strategies to inform and influence individual and community decisions that affect health.[5] It links the fields of communication and health and is increasingly recognized as a necessary element of efforts to improve personal and public health.
Dissemination	The active and targeted distribution of information and interventions to a specific public health or clinical practice audience via determined channels using planned strategies.[6,7] The intent is to spread knowledge and the associated evidence-based interventions in order to enhance the adoption and the implementation of the information and/or intervention.[8,9]
Adoption	The decision of an organization or a community to commit to and initiate an evidence-based intervention.[6,7]
Implementation	The use of strategies to integrate evidence-based health interventions and change practice patterns within specific settings.[6-8]

We deliberately avoid the term "translation" in our review because it has broad, and diverse, definitions. Rather, we focus on components of translation, specifically communication and dissemination, and on a special issue in communication, that of communicating uncertainty. Adoption and implementation processes to integrate evidence-based practices successfully into health care delivery to improve health outcomes are beyond the scope of this review.

Communication Strategies to Promote the Use of Health Care Evidence

Government agencies and institutions, advocacy groups, media organizations, researchers, and other interested stakeholders can all carry out communication activities. They use various strategies to communicate evidence so that target audiences can understand it better; the strategies are meant to increase the probability that recipients pay attention to the messages conveyed.[10,11] People are motivated to process information actively when they perceive it to be personally relevant. This attribute can reflect dimensions such as the number and magnitude of consequences relevant to them and the match of the information to an existing need.[12]

For purposes of our review, communication strategies fall into the broad area of "health communication" and focus on making evidence interpretable, persuasive, and actionable. The John M. Eisenberg Center for Clinical Decisions and Communications Science translates AHRQ's comparative effectiveness review information to create a variety of materials ranging from evidence summaries to decision aids and other products.

Overview of Four Main Communication Strategies

To our knowledge, no overarching framework of communication strategies exists to guide this part of our review. Multiple systematic reviews, however, have explicated key communication strategies that are of interest to the field. They include four core constructs:

1. **Tailoring the message—**Communication designed for an individual based on information from the individual.
2. **Targeting the message to audience segments—**Communication designed for subgroups based on group membership or characteristics such as age, sex, race,

cultural background, language, and other "psychographic" characteristics such as a person's attitudes about a particular subject matter.

3. **Using narratives—**Communication delivered in the form of a story, testimonial, or entertainment education.
4. **Framing the message—**Communication that conveys the same messages in alternate ways (e.g., what is gained or lost by taking an action or making a choice).

Table 2 summarizes recent evidence for the effectiveness of the four communication strategies that we examine: tailoring the message, targeting the message to audience segments, using narratives, and framing the message. These systematic reviews focused on the effectiveness of the communication strategies *relative to not using any strategy*, that is, relative to "usual practice." Thus, these reviews establish the contribution of each strategy compared with not using any communication strategy. By contrast, our focus is on the *comparative* effectiveness of different strategies.

Table 2. Systematic, meta-analytic, or theoretical reviews supporting various communication strategies

Author and Date Number in Study Search Dates	Communication Strategies	Main Conclusions Supporting Inclusion
Noar et al., 2007[13] N=58,454 Through 2005 Lustria et al., in press[14] N=20,180 1999-2009	Tailored communication	Tailored communication delivered via print or the Internet is more effective than nontailored communication in increasing knowledge and changing behavior. Effect sizes can vary based on length of followup, variables tailored, type of behavior, population studied (general vs. chronic illness), and number of intervention contacts.
Slater, 1995[15] Nonsystematic review Noar et al., 2009[16] N=94,896 1998-2007	Targeted communication to audience segments	Communication that is targeted to audience segments is a strategy used to make information more relevant based on group membership characteristics. Characteristics can be determined by role, demographic, or social-psychological variables. Although we have not found a systematic review on this approach, meta-analysis shows its practice is more common in large-scale communication efforts owing to its potential effectiveness.
Hinyard and Kreuter, 2007[17] Theoretical review N not reported Winterbottom et al., 2008[18] N=3,986	Narratives	Narrative forms of communication increase information processing and raise the persuasiveness of messages. People become transported into a situation that can enhance emotions, attitudes, and behaviors.
O'Keefe and Jensen, 2006[19] N=50,780 Through 2006 Latimer et al., 2010[20] N=6,679 Through July 2008	Message framing	Messages framed as emphasizing the benefits of preventive action are significantly better for influencing behavior than loss-framed messages, although the difference is small.

Tailoring

As with many other communication strategies, the rationale behind creating tailored communication is that it can maximize the relevance of the communication to its intended

audience. Rimer and Kreuter (2006)[21] argued that tailoring message content to an individual's informational needs and interests can elicit greater cognitive elaboration by increasing its perceived relevance.[22-24] Tailoring is a multi-step and multi-dimensional process that involves assessing an individual's characteristics, creating individualized messages, and then delivering these messages.[22,23] A typical tailoring study will first collect data from individuals regarding various psychosocial behavioral determinants. It will then use conceptually or empirically based algorithms—usually computer driven—to process each person's data and generate customized feedback to meet that individual's unique needs.

Targeting

Targeting (also referred to as audience segmentation) involves developing a single intervention approach for a defined population subgroup that takes into account characteristics that the group shares (e.g., age, sex, race, ethnicity, spoken language).[23] Tailored communication is intended to reach a specific individual; by contrast, targeted communication is intended to reach some population subgroup. Once those developing communications have segmented an audience or population in one (or more) ways, they should then design the messages to be maximally effective for that target subgroup (or subgroups). They can accomplish this by manipulating language, visuals, music, or choice of behavior topic. As with tailoring, message targeting is expected to enhance the perceived personal relevance of a message.

Narratives

Narrative messages are defined as "story-like prose pieces that focus on elaborating one example of an event, and they provide appealing detail, characters, and some plot, presented in either the first or third person."[18, p.2080] The characters and the situations in stories serve as models for emulation and learning. Some narratives include personal stories, case histories, anecdotes, and testimonies (e.g., a personal account of an individual's experience in donating an organ to a sibling). Evidence is mounting for the benefits of narrative health messages in promoting persuasion and behavior change.[17,18,25,26]

Framing

Appeals aimed at persuading individuals to perform healthy behaviors or avoid risky behaviors can be framed in different ways. Health messages are framed or presented within a specific context to promote or enhance comprehension. For instance, messages might emphasize the positive or negative aspects of a situation, commonly known as gain/loss framing. Gain- and loss-framed messages are factually equivalent. Previous research has found that gain-framed messages are significantly better than loss-framed messages.[19,20]

Interactions across Techniques and Generic Approaches

Communication techniques do not necessarily occur in isolation. One possible reason that prior literature often reports no or mixed effects from these four strategies[27] may be that message features moderate the effects of the strategies. For example, features such as the use of narratives may affect involvement and message relevance and intensify or minimize the effect of message framing on behavioral outcomes.

To date, most research fails to explore potential interactive effects of the strategies. However, because the content of a message typically contains several different features, most messages will likely combine various strategies (e.g., loss-framed narrative or gain-framed statistical

evidence). In addition, some techniques can be present in a study because of the nature of the variable itself. For example, with framing, every statement that connects a recommended action to some health outcome can be said to have either a positive or a negative frame. We will consider the potential interplay that may occur when messages contains multiple persuasive techniques by reporting any interaction effects observed in the primary studies.

Finally, several other communication approaches involve applying plain language principles or using theoretically driven messages. These approaches are widespread and can be considered best practices, but we excluded them from this review because they are general approaches used across many different communication techniques.

Dissemination Strategies to Promote the Use of Health Care Evidence

Dissemination is the active and targeted distribution of information or interventions via determined channels using planned strategies to a specific public health or clinical practice audience.[6,7,28] Dissemination has been characterized as a necessary but not sufficient antecedent of adoption and implementation. In contrast to diffusion, which is a passive, informal process, dissemination is a formal, planned process with the intent of spreading knowledge and associated evidence-based interventions to stimulate adoption and enhance the integration of the evidence, information, or intervention (or combinations of these) into routine practice.[6-9,29]

Dissemination strategies involve "packaging" the evidence, information, or intervention in different ways and using a variety of channels to reach the target audience(s) within or across geographic locations, practice settings, or social networks. Dissemination is often described as a "push/pull effort," with some strategies directed toward increasing the reach and accessibility of the evidence (i.e., push) and other strategies directed toward increasing the receptivity or readiness of the target audience (i.e., pull). This push/pull description has also been used more generally to describe an approach to closing the overall research-to-practice translation gap.[30-32]

Outside the United States, the phrases "knowledge translation" and more recently "knowledge exchange" are used to reflect iterative cycles of feedback and involvement of the target audience(s) in generating and incorporating evidence into routine practice. Furthermore, the concept of "knowledge brokering" describes the organized way in which the iterative process of knowledge exchange occurs, including dissemination.[7,33]

Evidence dissemination has several very broad goals relating to evidence and information: (1) to increase their reach to a variety of audiences; (2) to increase people's motivation to use and apply such information; and (3) to increase people's ability actually to use and apply evidence. A recent narrative review of dissemination and implementation research models identified 11 dissemination-only models and an additional 16 combined dissemination/implementation models with a predominant focus on dissemination.[28] In examining influences that help spread innovations along the continuum between passive diffusion of information and active dissemination, Greenhalgh et al. created an inventory of strategies that aim to influence individual, social, and other adopters and improve reach, motivation, or ability.[29]

Other systematic reviews and dissemination research show that active dissemination strategies are more effective than passive strategies.[34] For example, in a synthesis of 41 systematic reviews, Grimshaw and colleagues reported that active, multifaceted approaches were most effective for changing provider behavior.[35] Educational outreach, academic detailing, and the use of local opinion leaders are the most consistently effective interventions reported.

Interventions that are theory-based, that incorporate two or more distinct strategies (i.e., that are multicomponent), or that do both, are consistently more likely to work than single interventions.[36,37] Moreover, the Internet, technological platforms for social networking, and Web 2.0 applications all involve active steps, and users can create and interact with information in ways that give classic theories of dissemination a new twist.[38]

Evaluating strategies for disseminating evidence and information entails specifying interventions and desired outcomes (such as adoption of the disseminated information or intervention at the individual or organizational level). It also requires consideration of two other components: mediators, which are processes through which dissemination occurs, and moderators, which are factors influencing the speed and extent of dissemination.[6]

Explaining Uncertain Evidence

Uncertainty is inherent in health care and evidence about health care.[39] It stems from multiple sources, including imperfect knowledge about scientific evidence, patients' preferences and circumstances, and how to apply judgment in decisionmaking.[39-42] Uncertainty interferes with both patients' and physicians' ability to derive appropriate meaning about illness, diagnostic tests, treatments, and prognosis and to use this information in meaningful ways.

Further, the experience of uncertainty can sometimes create aversive psychological[43] responses. For example, the uncertainty in the 2009 U.S. Preventive Services Task Force (USPSTF) breast cancer screening recommendations for women ages 40 to 49 years created significant controversy and left some women more confused than helped.[44] Such confusion may sometimes lead end-users to avoid health evidence in an attempt to control anxiety or manage hope.[14,15]

In the context of evidence translation, uncertainty creates multiple challenges. These include difficulties in (1) determining whether preventive services and treatments should be implemented in clinical practice, (2) determining for whom and in what settings preventive services and treatments should be implemented, and (3) communicating evidence so that end-users can make informed decisions.

To date, the vast majority of work on communicating uncertainty has focused on the narrow realm of stochastic uncertainty: the likelihood or probability of an event occurring. This work has generally focused on alternate presentations of disease risk, side effects, treatment benefits, and treatment harms. Multiple systematic evidence reviews and randomized trials[45-49] have demonstrated that:

- Qualitative or non-numeric presentations of probability (e.g., "likely," "certain," "rare") are open to individual interpretation.[46,49]
- Percentages and "x/1,000" presentations are more understandable than "1 in x" presentations of probability;[49-51] "x/1,000" presentations are better than percentage presentations for representing conditional probabilities.
- Using the same denominator in "x/1,000" presentations[49,51,52] facilitates understanding.
- Absolute risk reduction and relative risk reduction are more understandable than number needed to treat presentations.[45-49]
- Absolute risk reduction tends to be less persuasive than relative risk reduction.[45-49]

Little research has focused on other concepts of uncertainty related to evidence translation. For our review, we synthesized the current literature to present our own framework of uncertainty as it relates evidence translation.[39-42,53-55] It encompasses the concepts of uncertainty

relative to the current scheme for grading the strength of evidence for AHRQ's Evidence-based Practice Center (EPC) program, including risk of bias, consistency, directness, and precision.[56] It also includes the broader ideas of net benefit (i.e. whether is more benefit than harm or vice versa) and applicability.[57,58] Detail about these domains follows in Table 3 below:

Table 3. Conceptual framework for uncertainty (KQ 3)

Component	Description
Overall strength of evidence	The strength of the evidence represents the degree of confidence that the estimates of effects are correct and represent the true effect. Strength of evidence grades are used to provide a comprehensive evaluation of the evidence and an assessment of whether additional evidence might change conclusions. Strength of evidence requires a value judgment based on the risk of bias, consistency, directness, and precision (see definitions below). When overall strength of evidence is high, uncertainty is low; when overall strength of evidence is insufficient or low, uncertainty is high.
Risk of bias	The risk of bias is the degree to which individual studies are protected from systematic errors or bias. Biases may result from study design, study conduct, or confounding by other external variables. Risk of bias is typically said to be low, medium, or high. This is analogous to ratings of the quality of the evidence, which are typically denoted good, fair, or poor.
Consistency	The consistency of a body of evidence reflects the degree to which studies present similar findings—in either direction of effect or magnitude of effect (or both). Evidence lacking consistency includes studies with greatly differing or conflicting effect estimates.
Directness	Directness is the degree to which the evidence either directly links the interventions to the outcome of interest or directly makes the comparison of interest.
Precision	Precision reflects the degree of random error surrounding an effect estimate with respect to a given outcome; such studies express dispersion around a point estimate of risk, such as a confidence interval, which indicates the reproducibility of the estimate.
Net benefit	Net benefit describes the balance or trade-offs in benefits and harms for prevention or treatment services. This is based on a judgment call by policymakers. Overall, evidence may reflect net benefit, clinical equipoise (benefit that is too close to call at the population level), or net harm. What constitutes a "sufficient" margin of benefit for evidence to provide "net benefit" is open to interpretation.
Applicability	Applicability reflects whether an intervention is expected to have the same effect in populations and settings where it was not studied, but might be applied.

Once the uncertainty in the strength of evidence and its component parts, net benefit, and applicability, is determined, various groups must communicate the information to end users. Explaining such findings and their implications in a way that enables end users to cope with uncertainty and process uncertainty is challenging.[59,60] Politi et al. suggest using subjective descriptions, various depictions of numbers, or visual aids to represent uncertainty and its degree,[53] but the best practices for communicating uncertainty are still unknown.[60,61] By optimizing the presentation of uncertainty, evidence creators, synthesizers, and disseminators can enhance awareness of the evidence, discussions around the evidence, and enable people to make the best possible decisions. This review seeks to compare techniques in communicating uncertainty related to evidence translation and assess their comparative effectiveness.

Scope and Key Questions

Scope of the Review

The purpose of this systematic review is to assist patients and consumers, clinicians and other providers, and purchasers and payers in making informed decisions that will improve health and health care at both the individual and population levels.[62] The Institute of Medicine's list of 100 priority topics highlights the importance of translating and disseminating findings from research evidence.[63] Many hope that better communication and dissemination of such research evidence will prompt wider and more effective use of the information.

Coupled with these mandates is the fact that the ad hoc Uncertainty Committee of the EHC Stakeholder Group is interested in promoting effective ways to communicate uncertainty about health and health care evidence to end users. The committee would like to know what approaches to conveying uncertainty increase the likelihood that audiences receiving such information will understand it and be able to factor it into their decisionmaking.

This systematic review has three related components; all focus on promoting informed decisions about health related behaviors and decisions among patients and providers. First, it addresses the comparative effectiveness of communicating evidence in various contents and formats that increase the likelihood that target audiences will both understand and use the information. Second, it examines the comparative effectiveness of a variety of approaches for disseminating evidence from those who develop it to those who are expected to use it. Third, it examines the comparative effectiveness of various ways of communicating uncertainty associated health-related evidence to different target audiences.

In this review, the interventions are communication strategies, dissemination strategies, and methods of explaining uncertainty. The strategies seek to influence health related behaviors and decisions. Due to the complexities of our questions, we present our work as three separate systematic reviews – one for communication, one for dissemination, and a third for uncertainty—each addressing a separate key question.

Key Questions

KQ 1

a. What is the comparative effectiveness of communication strategies to promote the use of health and health care evidence by patients and clinicians?
b. How does the comparative effectiveness of communication strategies vary by patients and clinicians?

KQ 2

a. What is the comparative effectiveness of dissemination strategies to promote the use of health and health care evidence for patients and clinicians?
b. How does the comparative effectiveness of dissemination strategies vary by patients and clinicians?

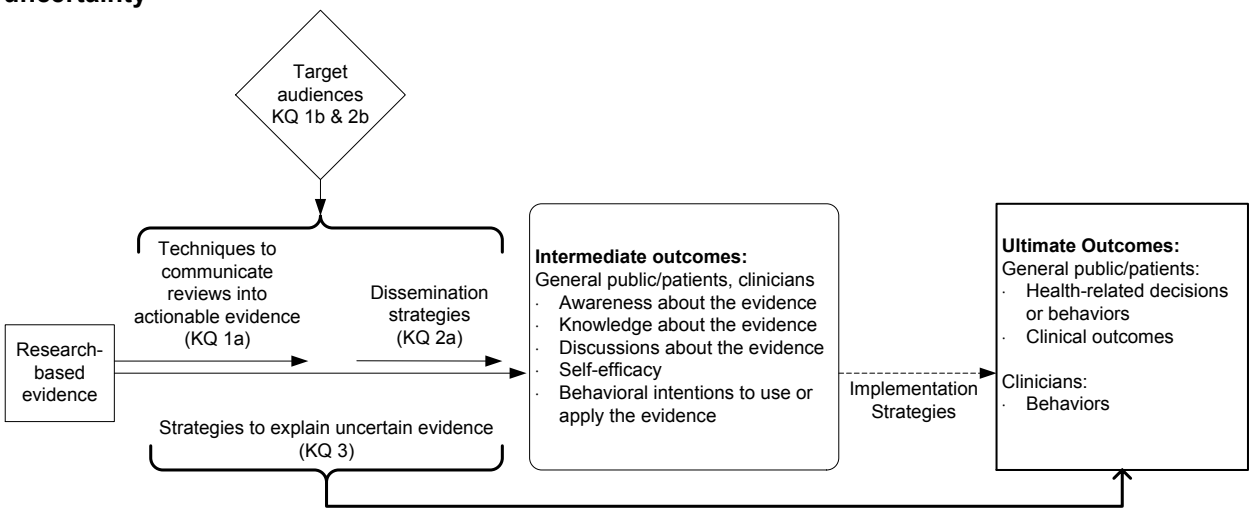
KQ 3

What is the comparative effectiveness of different ways of explaining uncertain health and health care evidence to patients and clinicians?

Analytic Framework

We present our analytic framework in Figure 1. As noted in the box to the far left, we examined studies that used research-based evidence as the source of information for their communication strategies (KQ 1) and dissemination strategies (KQ 2). For KQ 1 and 2, we specifically defined research-based evidence as evidence that has been assembled, reviewed, and presented by evidence developers and that has been used to make recommendations. For KQ 3, however, we accepted any type of evidence presented given the paucity of overall literature. (See Methods section for more specific inclusion and exclusion criteria for evidence.)

Figure 1. Analytic framework for communicating and disseminating strategies and explaining uncertainty



Strategies and techniques discussed in this review could be beneficial for several audiences. Such audiences include (1) patients and the general public and (2) clinical service providers, including physicians, nurses, midlevel providers, pharmacists, and others who deliver health care; in KQs 1a and 2a, we examine the effect of interventions in aggregate across these populations. Because the effects of interventions can differ for different target populations, we also examined (in KQ 1b and 2b)—how the effectiveness of communication and dissemination strategies vary across target audiences, including patients and clinicians. For KQ 3, we focused on studies that explored communication techniques to explain uncertain evidence.

We included studies that examined both intermediate and ultimate (distal) outcomes, as shown in in Figure 1. Intermediate outcomes can be awareness of the evidence, knowledge of the evidence, discussions about the evidence, self-efficacy (or one's confidence) to use the evidence, and behavioral intentions to use or apply the evidence. Ultimate outcomes include be the following: for patients—health-related decisions or behaviors and clinical outcomes; for clinicians—behaviors. We expected that most studies focused on intermediate outcomes because they occur sooner and, thus, are more practical to study.

Populations, Interventions, Comparators, Outcomes, Timeframes, and Settings Covered By the Key Questions

Below we describe the populations, interventions, comparators, outcomes, timeframes and settings (PICOTS) for our review (Table 4).

Table 4. Population, intervention, comparators, outcomes, and settings (PICOTS)

Domain	Description
Population	Recipients of health and health care evidence, also called "target audiences," which include: • Adult patients and the adult public at large • Clinicians, including physicians, nurses, mid-level providers, and/or pharmacists
Interventions	Specific clinical interventions, which include: **Strategies to communicate evidence, such as** • Tailoring the message • Targeting the message to audience segments • Using narratives • Framing the message • Using a multipronged approach with any of the communication techniques described above (e.g., tailoring and targeting) **Strategies to disseminate evidence, such as those that:** • Increase reach of the evidence (e.g., telephone; postal mail/e-mail; electronic/digital media, social media, mass media, interpersonal outreach) • Increase people's motivation to use and apply the evidence (e.g., opinion leaders, champions, social networks) • Increase people's ability to use and apply the evidence (e.g., additional resources, skills-building) • Use a multipronged approach with any of the dissemination strategies described above (e.g., social marketing, academic detailing) **Techniques to explain uncertain evidence, such as:** • Different presentation formats (e.g., graphical, numeric, non-numeric) • Any communication technique, including the ones above and hypothetical situations
Comparators	Alternate presentations of the specified interventions
Outcomes	Specific outcomes, which include: **Intermediate outcomes for all target audiences** • Awareness of the evidence • Knowledge about the evidence • Discussions about the evidence • Self-efficacy to use the evidence • Behavioral intentions to use or apply the evidence **Ultimate outcomes for patients** • Health-related decisions or behaviors • Clinical outcomes **Ultimate outcomes for clinicians** • Behaviors
Timing	Any length of followup is permissible
Settings	Clinical or community settings in the United States, such as • Inpatient and outpatient settings and clinics of all types • Academic health care institutions • Churches, settings for fraternal organizations, professional or social clubs, pharmacies • Homes

Organization of This Report

In the next sections, we describe the methods used in this review. We then present three separate results sections for KQ 1, KQ 2, and KQ 3, respectively. We then discuss our conclusions and the implications of our results, limitations of the evidence base and this review,

and important research gaps. Appendix A documents our search strategies. Appendix B lists all studies we reviewed at the full-text stage but excluded and the reason for exclusion. Appendix C contains the quality assessments (risk of bias) of the included studies. Appendices D, E and F contain the evidence tables for KQ1, KQ2 and KQ3 respectively.

Methods

The methods for this systematic review generally follow those of the *Methods Guide for Effectiveness and Comparative Effectiveness Reviews* for the Agency for Healthcare Research and Quality.[64] In this section, we explicate our topic refinement process and explain our literature search strategies (e.g., inclusion and exclusion criteria, search and retrieval process). We also describe methods of abstracting relevant information from the eligible included articles and our approach to data synthesis. We also discuss our criteria for rating the quality of individual studies and for grading the strength of the bodies of evidence for the major comparisons and outcomes of interest.

Topic Refinement and Review Protocol

To define the scope of our review and make it maximally responsive to stakeholders, such as guideline developers and policymakers, we engaged in a public process of development and refinement of key questions (KQs) for the review. Initially, we engaged a panel of experts in health communication, guideline development and implementation, and risk communication to solicit input on some KQs that our research team proposed. Using expert input, we then refined the KQs, and AHRQ posted them on their website for public comment on March 5, 2012 for 4 weeks. We then drafted a protocol and recruited members of a technical expert panel to provide high-level content and methods expertise throughout the review process. Our key informants and technical experts included representatives from the following disciplines: communication sciences, social marketing, health behavior, epidemiology, dissemination and implementation sciences and medicine.

Literature Search Strategy

In the Introduction, we set out the KQs in detail; Figure 1 provided the analytic framework that guided much of our work. As described below, we needed three sets of searches to cover the three main topics: (1) communication techniques to promote use of health-related and health care evidence and how they vary by patients and clinicians (KQ 1a and 1b); (2) dissemination strategies to promote use of health-related and health care evidence and how they vary by patients and clinicians (KQ 2a and 2b); and (3) different ways to explain uncertain evidence (KQ 3).

Search Strategy

We systematically searched, reviewed, and synthesized the scientific evidence for each KQ separately. Databases included MEDLINE®, the Cochrane Library, Cochrane Central Trials Registry, PsychInfo, and the Web of Science. We did not conduct additional searches for gray literature.

To identify articles relevant to each KQ, the EPC librarian began with three focused MEDLINE searches on the topics noted above, We used a variety of medical subject headings (MeSH terms) and major headings, free-text and title and abstract text-word searches (Table 5; Appendix B documents the exact search strings). Search results were limited to studies on humans published from 01/01/2000 to May 15, 2012 for communication and dissemination given the prior systematic reviews, and from 01/01/1966 onward to May 15, 2012 for uncertainty given the lack of prior reviews on this specific topic.

Table 5. Initial literature search terms for each of the targeted searches

Interventions	Search Terms for MEDLINE ®(PubMed)
KQ 1: Communication techniques to promote the use of health and health care evidence	"Information Dissemination/methods"[Majr] OR "Decision Making"[Majr] OR "Patient Education as Topic"[Mesh] OR "Narration"[Majr] OR OR] OR "Persuasive Communication"[Majr] OR "Health Education/methods"[Majr]
KQ 2: Dissemination strategies to promote the use of health and health care evidence	"Diffusion of Innovation"[Mesh] OR "Information Dissemination"[Mesh] OR "Evidence-Based Medicine/education"[Mesh] OR "Evidence-Based Medicine/methods"[Mesh] OR "Information Services/utilization"[Mesh] OR "Practice Guidelines as Topic/standards"[Mesh] OR "Guideline Adherence/statistics and numerical data"[Mesh] OR "] OR "Physician's Practice Patterns/standards"[Mesh] OR "Physician's Practice Patterns/statistics and numerical data"[Mesh] OR "Physician's Practice Patterns/trends"[Mesh] OR "Social Marketing"[Mesh] OR "social marketing"[tiab] OR "academic detailing"[tiab] OR "dissemination strategy"[tiab] OR "dissemination strategies"[tiab] OR (disseminat*[ti] AND guideline*[ti])
KQ 3: Methods of explaining uncertain health and health care evidence	("Uncertainty"[Mesh] OR uncertainty OR "low evidence" OR "conflicting evidence" OR "missing evidence" OR "strength of evidence" OR "Research Design/statistics and numerical data"[Mesh] OR "Therapeutic Equipoise"[Mesh] OR ambigu* OR complexity OR vagueness OR precision OR "risk of bias" OR "Bias (Epidemiology)"[Mesh] OR "net benefit") AND ("Communication"[Mesh])

Using analogous search terms, the librarian searched the Cochrane Library and Cochrane Central Trials Registry for trials on these topics. She searched PsychInfo for communication and uncertainty articles given the high likelihood of relevant publications in the psychological literature and the Web of Science to trace citations of known uncertainty frameworks and to capture articles on uncertainty.

To limit KQ 1 and KQ 2 searches to relevant *comparative* effectiveness literature, we further limited all searches to comparative effectiveness studies by including only studies that had any of the following keywords: comparative effectiveness, evidence based or evidence-based, and recommendation or recommendations. This process was analogous to limiting MEDLINE searches with text words. We did not further refine KQ 3 results given our broader approach to this literature

We expected some overlap in results across these searches. We removed duplications in our EndNote database and tracked the yield from each search.

We hand-searched bibliographies of included articles. In addition, in an effort to avoid retrieval bias, we manually searched the reference lists of landmark studies and background articles on this topic to look for any relevant citations that electronic searches might have missed.

Inclusion and Exclusion Criteria: Overall

Criteria for inclusion and exclusion of studies address both the PICOTS model (population, interventions, comparators, outcomes, timeframes, and settings; see Introduction) and other important study design and publication issues. Table 6 presents the inclusion/exclusion criteria common to all three KQs; Table 7 defines the inclusion criteria applied to admissible research evidence for KQ 1 and KQ 2. We present other KQ-specific inclusion/exclusion criteria in Tables 8, 9, and 10, respectively.

Table 6. General inclusion/exclusion criteria for all key questions

Category	Criteria for Inclusion	Criteria for Exclusion
Language	English	All non-English publications
Dates of publication[a]	• 01/01/2000 to May 15, 2012 for communication and dissemination • 01/01/1966 to May 15, 2012 for uncertainty	• Anything published through 12/31/1999 for communication and dissemination • Anything published through 12/31/1965 for uncertainty
Study design	• Individual randomized controlled trials • Cluster randomized controlled trials • Quasi-experimental trials (KQ 3 only) • Nonrandomized trials (KQ 3 only)	• All nonexperimental studies • Qualitative research
Publications	Complete articles	• Systematic reviews[b] • Meta-analyses • Protocols • Studies published only as abstracts • Studies with no original data (i.e., no experimental data) • Narrative reviews • Editorials, letters to editors, and similar publications
Study Populations	Adults (≥19 years) • General public and patients • Clinicians	• Children (<19 years) • Incarcerated populations • Federal and State policymakers
Interventions	• Techniques and strategies as specified for individual key questions	• See Tables 8-10 and associated text.
Comparators	• Alternate presentations of specified interventions for individual key questions	• Comparisons with usual practice (except for KQ 3 when the evidence is sparse)
Outcomes	• Prevention and treatment	End of life
Timeframes	• No limits to study duration	
Settings	• Inpatient and outpatient settings and clinics of all types • Academic health care institutions • Community-based settings such as churches, fraternal organizations, professional or social clubs, pharmacies, homes	• Primary and secondary schools • Prisons and jails
Geographic locations	• Austria, Australia, Belgium, Brazil, Canada, Denmark, Finland, France, Germany, Greece, Ireland, Israel, Italy, Japan ,Luxembourg, Netherlands, Norway, Poland, Portugal, South Africa, Spain, Sweden, Switzerland, Turkey, United Kingdom, United States	• Any other country not specified for inclusion
Sample sizes	• N ≥ 100 total individuals in the study • No limits on size of clusters	• N < 100 total individuals in the study
Other	Access to entire article	Inability to retrieve full article

[a]We will update searches when the draft report is out for peer review.

b We completed a hand-search of systematic reviews and use systematic reviews only for background information.

Abbreviations: KQ = key question; N = number

A few specific decisions we made regarding inclusion and exclusion criteria bear special mention. First, to improve the overall quality of included findings, we focused on randomized trials with at least 100 total individuals in the study (e.g., 50 individuals per arm in a study with

two arms) to prioritize studies with greater statistical power and less chance for confounding. Second, we limited our review to evidence communicated and disseminated to clinicians (a category that included physicians, nurses, midlevel providers, and pharmacists) and patients. Third, we limited our evidence to numerous countries specified in a recent analysis of the world systems of nations that are likely to have evidence relevant for our target audiences.[65]

Finally, for communication and dissemination (KQ 1 and KQ 2), we searched only from 2000 to the present for two main reasons: (1) comparative effectiveness work became more common after 2000 and (2) multiple systematic reviews on communication and dissemination appeared after 2000, thus assuring that we could capture relevant older literature through those publications. We also we limited our review to studies communicating or disseminating research evidence that had been a systematically assembled, reviewed, presented, or used to make recommendations about clinical practice. Table 7 documents these criteria. By applying these criteria, we excluded studies communicating or disseminating evidence developed or assembled through a consensus process or by individual researchers in the process of developing or testing interventions. Doing so allowed us to compile a clearly definable set of research evidence and guidelines to be communicated and disseminated to end users.

Table 7. Inclusion and exclusion criteria for research evidence to be communicated or disseminated

Inclusion Criteria	Exclusion Criteria
Systematic review evidence or guidelines generated by governmental organizations or agencies, such as: United States Preventive Services Task Force (USPSTF); the Community Preventive Services Task Force; Cochrane Collaboration; National Institute for Health and Clinical Evidence (NICE); National Institutes for Health (NIH) agencies (such as National Health Lung and Blood Institute (NHLBI), National Institute of Diabetes and Digestive and Kidney Disorders (NIDDK), National Cancer Institute (NCI); AHRQ-funded Evidence-based Practice Centers (EPCs).	Evidence based solely on a consensus process.
Systematic review evidence or guidelines from a professional organization or society, such as American Medical Association (AMA) or the American Cancer Society (ACS) if information in the article indicated an attempt to base on evidence (literature review and not just consensus).	Evidence based on a single intervention trial or on prior work by the “evidence” authors.
Evidence from guidelines in the National Guideline Clearinghouse (NGC) because the NGC expects recommendations to be based on authoritative sources using systematic review techniques.	Evidence for which a specific guideline or systematic review from a recognized body was not cited as the source.
Evidence from government-supported (e.g., NCI) research consortiums (such as the NCI Breast Cancer Research Consortium) that is compiled using acceptable methods such as systematic reviews and/or meta-analyses and that presents guidelines or recommendations to inform practice or behavior change (or both).	Evidence based on conference proceedings.
	Evidence based on a compilation or combination of multiple sources and heavily adapted. These combinations of multiple sources were often compiled using consensus-based decisions and nonsystematic procedures.

Inclusion and Exclusion Criteria Specific to Communication Techniques

For KQ 1, strategies of interest include tailored communication, communication targeted at audience segments; use of narratives; and message framing (Table 8). These strategies are designed to make information clearer, easier to understand, and more relevant to end-users. We included studies that compared two or more of the included communication strategies head to head (e.g., tailoring versus targeting).

Table 8. Included communication strategies and approaches for key question 1

Type of Communication Strategy	Included Approaches to Communication
Tailoring the message:	• Using a aomputerized database of messages that can be combined in response to answers to preprogrammed questions asked of an individual • Applying an electronic algorithm to design messages based on individual input regarding a limited number of questions • Trying to direct messages to individuals' status on key theoretical determinants (knowledge, outcome expectations, normative beliefs, efficacy. or skills) of the behavior of interest • Incorporating recognizable aspects of participants to convey (implicitly or explicitly) that the messages are designed specifically for them. This is more than a personalized letter (e.g., "Dear Jane") • Providing messages to participants about their psychological or behavioral states. Individualized feedback may be provided synchronously (e.g., via a chat function, telephone, or face-to-face) or asynchronously (e.g., via e-mail or discussion board, or mail)
Targeting the message	• Manipulating language, visuals, music, or choice of behavior topic in ways that make the message more interesting, relevant, or appealing to specific subgroups.
Using narratives	• Invoking personal stories, case studies, anecdotes, testimonials, experiential sharing (e.g., personal account of an individual's experience in donating an organ to a sibling) • Using entertainment education (e.g., talking about issue in a soap opera storyline) or photo novellas or graphic novels

Framing the message:	• Creating messages that emphasize the positive consequences of compliance are referred to as *positive (gain) frame*, whereas those that stress the negative consequence of noncompliance are denoted as *negative (loss) frame*. Studies should explicitly state that the stimuli differed in terms of gain or loss frame. For example,
	  o Positive (gain) frame: "Get active! Enhance your health!" vs. "A lack of activity increases risk for diabetes."
	  o Negative (loss) frame: "With drug X, you have a 5% chance of dying" vs. "With drug X, you have a 95% chance of surviving."

More than one of the above strategies	• A multi-component approach uses several communication strategiesin concurrent combination or in sequence to increase understanding of the evidence or information. • Multicomponent interventions are important to this review only to the extent that they are compared with another intervention that is different by only 1 or more aspects

Multi-component approaches seek to increase the overall impact of the information across geographic and practice settings and target audiences; they also aim to raise recipients' understanding of the information. A multi-component approach uses several (or at least two) communication strategies either simultaneously or in sequence. We included all studies that used a multi-component approach that had a combination of two or more communication strategies (e.g., tailoring and targeting) compared with a single strategy.

We excluded studies that compared one of these communication strategies with "usual practice" (i.e., steps that are essentially standard procedures and do not represent any included

strategies that serve as interventions of interest). Several prior reviews have previously examined communication techniques against only usual practice. We also excluded studies that compared permutations of the included communication strategies, for example, comparison of two different ways of using narratives. We excluded studies that examined interpersonal communication techniques because these tend to be more costly to implement and less practical for reaching large audiences. Finally, given the volume of other research (e.g., from the Cochrane Collaboration) focusing on decision aids, we included studies of decision aids only when they were based on evidence-based guidelines and met the other inclusion criteria above. To be included, studies must have used a decision aid as a communication strategy or dissemination technique.

Inclusion and Exclusion Criteria Specific to Dissemination Strategies

For KQ 2, we focused on *active dissemination strategies* that involve efforts to spread evidence-based information via specific strategies and channels. We included active dissemination strategies that are designed to do one or more of the following (see Table 9): (1) increase the reach of information (e.g., postal and electronic mail; electronic/digital, social, and mass media); (2) increase people's *motivation* to use and apply evidence (e.g., using champions, opinion/thought leaders, peer and social networks); and (3) increase people's *ability* to use and apply evidence (e.g., by also providing additional resources or information; skills-building efforts).

We sought to evaluate the comparative effectiveness of "pull" strategies that moderate the receptivity of the target audience(s) to the disseminated information. For example, push strategies include the use of traditional dissemination outlets (mail, print, or broadcast media) but may also include newer media such as Web or mobile applications, social networking sites, or text messaging. Examples of organizational-level pull strategies include packaging the information or the intervention (or both) so that characteristics of the intervention that influence organizational adoption (e.g., relative advantage, compatibility, and complexity[31]) are easy to find. Individual-level pull strategies include packaging, distributing, or marketing information or interventions in ways that increase recipient's motivation, skills, or capacity to use the evidence. For example, deploying toolkit-type materials for patients and providers and providing access to training sessions for the specific skills needed to implement the information or intervention are often used.

We included head-to-head comparisons (e.g., increase reach using social media vs. increase motivation using peer networks), comparisons of different strategies with the same broad aims (e.g., increase reach using social media vs. increase reach using digital media), or comparisons of the same strategy (e.g., one type of peer network compared with another). Multi-component strategies use several dissemination strategies in concurrent combination or in sequence to increase the reach of evidence and to enhance end users' motivation to adopt, use, or apply the evidence. We included all studies that use a multi-component approach with a combination of two or more dissemination strategies (e.g., social marketing, academic detailing) compared to single strategies.

Table 9. Included dissemination strategies for key question 2

Type of Dissemination Strategy	Included Approaches to Dissemination
Improve reach of evidence: Distributing evidence widely to many audiences and across many settings extends the numbers and types of recipients.	• Postal: Any information delivered to a new destination via a human carrier employed by a government-affiliated postal service or a for-profit mail or parcel delivery service such as FedEx™ or UPS® • Electronic and digital media: Any information delivered via telephone or web-based email, text messages, or electronic programs such as personal digital assistant (PDA) resources or phone apps • Social media: Any information delivered via Internet-based social networking sites such as Facebook, Twitter, YouTube, myspace, foursquare, and LinkedIn. Sometimes problem- or group-specific social networks exist for professional organizations or patient subgroups; these would fall into social media as long as they have a "social" network component as described above. • Mass media: Any information delivered via television, radio, print newspapers, print magazines, or billboards • Interpersonal verbal group or individual outreach: Information delivered via telephone, webinar, or in-person visits, including purposeful delivery of brochures or pamphlets, but without any motivational component. The audiences can include: pharmacists, nurses, doctors, counselors, or other clinicians.

Motivate recipients to use and apply evidence: Using a variety of authoritative experts or spokespersons to increasing interest in or acceptability of the evidence or related recommendations may promote enthusiasm or action on the part of clinicians or patients.	• Champions (cheerleaders): People who take ownership of the evidence and visibly promotes it within their own organization or across other settings. Champions help overcome social and political pressures imposed by an organization, provite a role model for personal commitment to the program, and involve others in its use.
	  o For example, an evidence developer might train or enlist the help of a local champion to promote evidence within his or her organization.
	• Opinion or thought leaders (frequently has an endorsing or persuasive element): Recognized expertswho lend their name to dissemination efforts to endorse the idea being disseminated and to establish credibility. They may or may not actually participate in the work and do not necessarily have any relationship with the organization to which evidence is to be disseminated. They could endorse the intervention, have a role in its development, or advise on strategies.
	  o For example, an opinion leader might be the CEO or the head of a department, an external expert in a particular field applicable to the evidence, or a well-recognized figure such as the US Surgeon General.
	• Social networks: A network of individuals who have a common perspective, relationships, or similar connection. The relationships can be informal (friends, peers, family) or formal (patient provider, nurses), but network members have defined role obligations. Peer networks provide a central and trusted source for information and might use multiple other dissemination strategies themselves (such as newsletters, journals, phone- and internet-based distribution, face-to-face conferences, peer-to-peer conversations, etc.)

Enhance recipients' ability to use and apply evidence (regardless of delivery mode) Providing additional resources about evidence or recommendations based on evidence, such as how they can be incorporated into current practice, or giving specific suggestions for change enhances a traditional dissemination strategy.	• Provision of supporting "how-to" materials: Includes physical materials that a health care practice might use to apply evidence in their activities. These might include giving tracking sheets to patients or giving risk calculators to clinicians. These might also include tailored toolkits that explain how to implement evidence-based recommendations from in specific settings. • Supporting materials *do not* include brochures, counseling resources, or resources that originate from the practice. They must originate from the evidence developer and be given to the end user. • Skill training, capacity building, and problem solving: Training in any skill that would allow appropriate use of evidence (to overcome barriers); might include training in recognizing the quality of evidence or the circumstances under which it can be reasonably used; also includes training in various counseling techniques that would facilitate evidence implementation and interactive seminars.

Table 9. Included dissemination strategies for key question 2 (continued)

Type of Dissemination Strategy	Included Approaches to Dissemination
More than one of the above strategies: Combining multiple dissemination strategies, including ways to increase reach, motivation, or ability, may be more effective than single strategies.	• A multi-component approach uses several dissemination strategies in concurrent combination or in sequence to increase the reach of evidence, enhance the end users' motivation to adopt and use or apply evidence. Multi-component interventions are important to this review only to the extent that they are compared with another intervention that is different by at least one otheraspect.

We excluded studies that compared the above strategies to "usual practice." In this case, this means passive, uncontrolled spread of information of evidence or no direct effort to spread information such as posting information to an evidence developer's website or posting scientific publications in a searchable database. The basic rationale is that passive dissemination strategies are generally not effective.[35] We also excluded studies that compared enhanced versions of the same strategy (e.g., monthly telephone calls vs. weekly telephone calls).

When investigators did not describe what a control group involved, other than to describe it as "usual practice," we excluded the study. In some instances, the authors may have said that the control strategy was "usual practice;" upon examination, however, we reclassified it as an active comparator. For example, a study might have described mailing a guideline as usual practice (or usual care), but for the purposes of this study, we considered that step to be the active strategy of "improving reach."

We excluded studies in which the primary purpose of the intervention was implementation (see definition in the Introduction), even when the intervention seemingly raised awareness or educated patients or clinicians (such as reminders at the point of care or audit-and-feedback). An example of implementation is when a clinical practice adopts and tries out a new treatment approach that is based on newly available health or health care evidence. Thus, if investigators were exploring how clinicians put a communication or dissemination approach into practice and were evaluating what impact that on their patients and patients' outcomes, then we considered that study (or that part of the study) to be implementation and either did not include the study (or omitted the findings for the implementation portion of the study).

Inclusion and Exclusion Criteria for Studies to Present Uncertain Evidence

Health-related and health care evidence inherently involves some degree of uncertainty. We focused this review on uncertainty in a body of evidence and how to communicate this uncertainty effectively to target audiences in ways that allow informed decisions. Given the early state of the literature on communicating uncertainty about evidence, our search for such studies was intentionally broad (i.e., inclusive) within the overall inclusion and exclusion criteria outlined above.

As defined in Table 3 in the Introduction, we examined studies that compared ways to explain seven types of uncertainty. Five come from the EPC program approach to grading strength of evidence: the overall grade for strength of evidence and the four principal domains used in deriving that grade—risk of bias, consistency, precision, and directness. We also considered studies that compared ways to explain net benefit (of either prevention or therapeutic services). Finally, we looked at the issue of applicability (i.e., generalizability or what is sometimes termed external validity). In this context, we looked at studies that attempted to

explain that although research evidence may exist on a particular topic, it may not be generalizable for one or more reasons.

Because our interest was in communicating uncertainty, we included any communication strategy that investigators used to communicate uncertainty. These could include non-numeric, numeric, or visual presentations of uncertainty or presentations using any of the communication techniques included for KQ 1 as shown in Table 10 below.

Table 10. Included communication strategies and approaches for key question 3

Type of Communication Strategy	Included Approaches to Communication
Non-numeric presentations	• Using words or sentences to describe the presence, degree, or meaning of uncertainty in medical evidence
Numeric presentations	• Using numbers to describe the presence, degree, or meaning of uncertainty in medical evidence
Visual presentations	• Using graphs, images, or figures to describe the presence, degree, or meaning of uncertainty in medical evidence
Tailoring presentation:	• Using messages that are personalized based on an individual's unique psychological characteristics (e.g. ambiguity aversion, optimism) that might affect their interpretation of evidence
Targeting presentation	• Manipulating the presentation of uncertainty to make it more interesting, relevant, or appealing to a specific subgroup of individuals
Narrative presentation	• Invoking personal stories, case studies, anecdotes, or testimonials to help individuals understand the presence, degree, or meaning of uncertainty related to medical evidence
Framed Presentation:	• Creating messages that present uncertainty in alternate contexts (e.g. relative to other more or less uncertain services) • Creating messages that present alternate consequences of uncertainty (e.g. "chances may be as high as" or "chances may be as low as")
More than one of the above strategies	• A multi-component approach uses several communication strategiesin concurrent combination or in sequence to increase understanding of the evidence or information. • Multicomponent interventions are important to this review only to the extent that they are compared with another intervention that is different by only 1 or more aspects

Unlike KQ 1 and KQ 2, we did not require that studies included for KQ 3 communicate uncertainty related to systematic reviews or guideline evidence. Instead, because of the overall paucity of evidence, we included studies communicating any type of evidence in either real world settings or hypothetical examples.

The following topics, although important, were beyond the scope of this review. We did not examine interventions designed to help individuals cope with uncertainty. We also excluded studies that compared alternative presentations of point estimates, as previous reviews on risk communication have well summarized these studies.[66-69] Finally, because our focus was on alternate ways of communicating uncertainty related to well-synthesized medical evidence that would assist with decision making about health alternatives, we excluded studies that addressed uncertainty related to: multiple causes of illness, changes in risks over time, lack of an individual's knowledge about evidence that is available, unclear values, trade-offs in care prompted by limited-resource settings, concerns about clinicians' competence, concerns about how a medical illness will affect family and friends, imperfect diagnostic testing, and uncertain prognosis.

Study Selection

Two trained members of the research team independently reviewed all titles and abstracts identified through searches for eligibility in terms of the overall or KQ-specific inclusion/exclusion criteria. Studies marked for possible inclusion by either reviewer underwent a full-text review. For studies without adequate information to determine inclusion or exclusion, we retrieved the full text and then made the determination. We tracked all results in an Excel database.

We retrieved and reviewed the full text of all articles included during the title/abstract review phase. Again, two trained members of the research team independently reviewed each full-text article for inclusion or exclusion on the basis of the eligibility criteria described earlier. If both reviewers agreed that a study did not meet the eligibility criteria, we excluded it. If the reviewers disagreed, conflicts were resolved by discussion and consensus or by consulting a third, senior member of the review team. All results were tracked in an EndNote database. Appendix B lists all studies excluded at this stage and the main reasons for exclusion. The disposition of all items (starting with the initial yields of the searches) through to articles finally retained for synthesis, are reported in a flow diagram conforming to Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) standards (see Figure 2 in the Results for KQ 1 section).[70] We accounted for studies reported in multiple articles.

Data Extraction

For studies that met inclusion criteria, trained reviewers extracted relevant information into specifically designed abstraction spreadsheets to facilitate the capture of all pertinent information from each article, including study design, characteristics of study populations, interventions, comparators, outcomes, settings, and results (Table 11). A second member of the team reviewed all abstractions against the accompanying article(s) for completeness and accuracy. Final approved abstraction spreadsheets were compiled and presented as evidence tables, which can be found in Appendix D, E, and F. These evidence tables formed the basis for the summary tables presented in the results sections to supplement text about synthesis of the evidence.

We relied on the analysis and comparisons provided by the authors. However, the review team had to calculate differences between groups (e.g., in mean values on a scale or percentages). We did this by subtracting the value for the intervention arm thought to have, or originally hypothesized to have, greater effect from the one expected to have a weaker effect. Because numerous findings led to negative differences (because of the original choices about the directionality of comparison), for KQ 1 the table indicates whether the difference was negative or positive and notes which group the findings favored. By favored, we mean which study group had the better result, namely a positive health behavior. For KQ 2 articles, the study authors sometimes did not make any predictions in terms of which arm should have the greater effect, therefore we report the absolute difference that emerged. These detailed findings are shown in Appendix E.

Table 11. Data items extracted

Data Extracted	Examples of Data Items
Study characteristics and methods	• Study design • Study objectives • Funding source • Geographic location • Sampling strategy • Eligibility criteria • Sample size • Units and methods of randomization • Measurement intervals • Sample attrition • Statistical analysis, including adjustment for multiple comparisons, clustering, and use of intention-to-treat analysis • Covariates used in the analysis
Population characteristics	• Age group • Gender or sex • Education • Race and/or ethnicity • Income • Insurance status • Health literacy or numeracy
Intervention and comparators	• Source of information • Clinical focus of evidence • Intervention format and delivery agent • Format for evidence presentation • Use of theory
Outcomes	• Definition of outcomes • Measures used • Source of outcome data
Settings	• Setting type • Descriptive characteristics
Results	• Results in intervention and control groups • Differences in effect between intervention and control group

Risk of Bias Assessment of Individual Studies

To assess the risk of bias of individual studies, we used criteria described in the AHRQ *Methods Guide for Effectiveness and Comparative Effectiveness Reviews*.[64] We used questions adapted from the RTI Item Bank,[71] the Cochrane Risk of Bias tool,[72] and prior work by the US Preventive Services Task Force.[57] We assessed the potential for selection bias (including attrition bias), measurement bias (such as performance bias, detection bias), confounding, and power. We also assessed potential biases in reporting.

We qualitatively synthesized the results to determine a rating of low, medium, or high risk of bias. In general, a study with a low risk of bias had a strong design, measured outcomes appropriately, used appropriate statistical and analytical methods, reported low attrition and little or no differential attrition, and reported methods and outcomes completely. Studies with a medium risk of bias were those with some bias, but not enough to invalidate results, and did not meet all criteria required for low risk of bias. These studies may have had some flaws in design or execution (e.g., imbalanced recruitment, high attrition) but they provided information (say, through sensitivity analysis) that enabled the reader to determine whether those flaws were not likely to cause major bias. Missing information often led to ratings of medium rather than low

risk of bias. Studies with a high risk of bias were those with at least one major flaw that was likely to cause significant bias and thus might have invalidated the results. Major flaws preclude the ability to draw causal inferences between the intervention and the outcome.

Two reviewers independently assessed the risk of bias for each study (see Appendix C for the final criteria we used and results). They resolved disagreements by discussion and consensus or by consulting a third, senior member of the team.

Data Synthesis: Overall

Studies included in our review compared a wide range of interventions and plethora of outcomes; they were sufficiently heterogeneous to preclude meta-analysis. Thus, we synthesized the data qualitatively by KQ. We paid particular attention to moderators of study effects as a way to explain any seemingly disparate findings. Possible moderators of interest for all key questions included risk of bias, study size, and target audience. We did not retain studies of high risk of bias for analysis, presentation in the results sections or strength of evidence grading.

Data Synthesis: Methods Specific to Key Questions

As noted in the introduction, we organized our report into three separate results sections specific to a KQ: communication, dissemination, or uncertainty. *Within* each section, we organized our results by the types of intervention strategies compared and then by outcomes, if possible.

For each subset of studies, we summarized key findings, including results in the experimental or quasi-experimental and the comparator groups and absolute differences between groups. If investigators did not report absolute differences between groups, we recorded the effect size that authors had reported and calculated an absolute difference. This approach gave us the best clinical interpretation of data.

Other than the overall moderators of interest noted earlier, we also looked at moderators specific to the KQ. These included:

- Communication techniques
  - Health literacy/numeracy level of audience
  - Intervention intensity or complexity (or both)
  - Message delivery setting
  - Message source
- Dissemination strategies
  - Care delivery setting
  - Type of media, mode, or channel (e.g., intervention format and delivery agent)
- Techniques for communicating uncertainty
  - Health literacy or numeracy of audience
  - Format of presentation (graphic, numeric, non-numeric, combination).

Strength of the Body of Evidence

We graded the strength of evidence on the basis of guidance established for the EPC program.[64,73] The EPC approach incorporates four required domains: risk of bias (including study design and aggregate quality), consistency, directness, and precision of the evidence. Table 12 defines the four overall grades for bodies of evidence that can be assigned. Grades reflect the

confidence that we have in the ability of the evidence to answer the KQs on the comparative effectiveness of the interventions in this review.

Table 12. Definitions of the grades of overall strength of evidence

Grade	Definition
High	High confidence that the evidence reflects the true effect. Further research is very unlikely to change our confidence in the estimate of effect.
Moderate	Moderate confidence that the evidence reflects the true effect. Further research may change our confidence in the estimate of the effect and may change the estimate.
Low	Low confidence that the evidence reflects the true effect. Further research is likely to change our confidence in the estimate of the effect and is likely to change the estimate.
Insufficient	Evidence either is unavailable or does not permit estimation of an effect.

Source: Owens et al., 2010[73]

Two reviewers independently rated the four domains for each intervention for each key outcome (listed in the analytic framework depicted in Figure 1); conflicts were resolved by group consensus. Two reviewers also independently derived the overall strength of evidence grade (resolving conflicts in the same way).

We adopted some conventions for assigning overall grades. First, when no studies were available or studies provided conflicting results, we graded evidence as insufficient. Second, when we had a single study, we graded evidence as low given that it was impossible to assess the consistency of evidence across settings and results would very likely change with additional testing.

For judging precision, we judged it as precise if: (1) confidence intervals were available, were reasonably narrow and did not cross minimal clinically important differences or the null; or if (2) confidence interval were not available, but the sample size was 400, which is a relatively conservative number. We judged it as imprecise if: (1) the confidence intervals crossed minimal clinically important differences or the null, or if (2) the sample size was less than 400.

To judge the strength of the evidence based on a single study, we applied the following criteria: (1) if imprecise and the risk of bias was moderate, we determined that the SOE was insufficient; (2) if precise and the risk of bias was moderate, we determined that the SOE was low; (3) if precise and risk of bias was low, we determined that the SOE was low (but discussed the issue as a team if the study was extremely large and across multiple sites, allowing consistency to be determined).

We present a summary of the strength of evidence for each intervention in the results section. Detailed strength of evidence tables can be found in Appendixes D, E, and F.

Applicability

We examined the applicability of the body of evidence for specific KQs by looking at characteristics that may limit applicability based on the PICOTS structure.[64,74] Such conditions may be associated with heterogeneity of treatment effect and the ability to generalize the effectiveness of an intervention to use in everyday practice. Examples of issues that may limit applicability include the following:

- Population: narrow eligibility criteria,
- Outcomes: different preventive behaviors or clinical conditions

- Settings: restrictions to certain types of health care or other institutions when the communication or dissemination activities might be carried out in many different locales or venues, and
- Timing: studies of different durations or points of followup that may have various implications for applicability.

Peer Review and Public Commentary

To be completed after peer review.

Results: Key Question 1: Communication Strategies

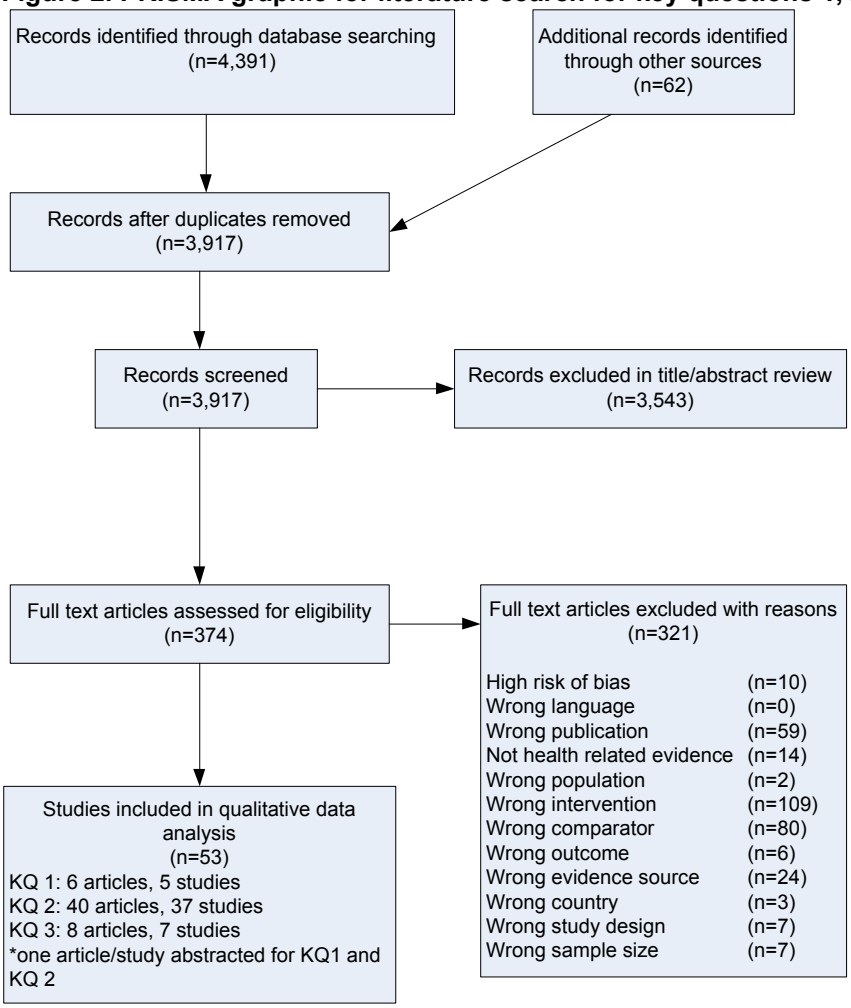
Results of Literature Searches for Key Questions 1, 2 and 3

We identified 3,917 articles from all sources (after removing duplicates) for all three key questions (see Figure 2). Two independent reviewers examined each abstract and applied our inclusion and exclusion criteria. Based on that process, we retained 374 articles for full text review. The majority of the full-text articles were classified to one or more key question (KQ) – 93 articles pertained to KQ1; 153 articles pertained to KQ2 (one that overlapped with KQ1); 80 articles pertained to KQ3, and 49 articles were classified as overlapping. Each overlapping article potentially applied to two or more KQs and was not classified into one KQ category.

Of the full-text articles, we excluded 321, leaving 53 articles for data abstraction. Six articles (representing five studies) are relevant to KQ 1; 40 articles (representing 37 studies) are relevant to KQ 2; and eight articles (representing seven studies) are relevant to KQ 3. Appendix B lists all articles excluded at the full-test review stage and the reason for exclusion.

This section presents the results for KQ 1: the effect of various communication strategies on both intermediate and distal outcomes. For KQs 2 and 3, we provide more information in the two sections that follow.

Figure 2. PRISMA graphic for literature search for key questions 1, 2 and 3



Introduction

The analytic framework for KQ1 appeared in the Introduction. For this KQ, we examined comparative studies of the following communication strategies: tailoring messages to individuals; targeting messages to audience segments; using narratives to convey messages; and using framing to convey messages to various end-users. For this KQ, as we noted in our methods, we included only randomized controlled trials (RCTs).

Some trials compared two strategies directly with each other (e.g., targeting vs. tailoring); others used a combination of strategies (e.g., targeting and tailoring vs. tailoring). The tables below describe individual trials and their results and document our SOE grades. Detailed evidence tables for KQ1 studies are in Appendixes D. As noted above, we retained six articles after full-text review that met inclusion criteria,[75-80] which report on five unique trials about communication strategies.

Description of Included Studies

Of the five included trials, we assessed one as low risk of bias[78]and four as moderate risk of bias.[76-78,80] None was assessed as high risk of bias. Four of the trials reported on the effects of intervention messages and materials on adherence to guidelines about breast or cervical cancer screening. One trial reported on the effects of the intervention materials on dietary behaviors. Because of the diversity of communication strategies and outcomes, we graded the SOE for three of the studies separately. We graded the SOE for two studies together because they examined targeting versus tailoring even though they addressed different outcomes.

The investigators tested these interventions in study populations in the United States. Sample sizes ranged from 174 participants to 5,500 participants. Four of the five trials used convenience samples.

Key Question 1: Communicating Evidence to Patients and Clinicians

Key Points

- **Targeting versus tailoring:** Findings were either counterintuitive or non-significant for the two studies that compared targeting with tailoring. In both studies, investigators hypothesized that the tailored version of the intervention would have a greater effect on screening rates than the targeted version. However, the *targeted* version was associated with greater likelihood of self-reported screening relative to the *tailored* version in one study. The authors attributed this unexpected finding to a possible "boomerang" effect or possibly a lack of meaningful difference between the two versions of the intervention. In another trial, there were no significant differences in dietary outcomes between those receiving the targeted or tailored version of the intervention. Adding a lay health worker to the tailored version of the intervention did not improve dietary outcomes (two trials; low SOE).
- **Targeting and tailoring versus targeting only**. Investigators found no statistically significant differences when they targeted an intervention to the subpopulation *and* personally tailored it to each study participant compared to a version of the intervention that was only targeted. They attributed the lack of differential impact to a possible

'ceiling effect' in the study population given the fairly high baseline screening rates (of about 80 percent) (one trial; low SOE).

- **Framing the message:** Loss framed messages appeared to have a more persuasive impact than gain framed messages in the two screening trials that used them. However, investigators used framing in conjunction with narratives (i.e., presenting the evidence using a story format versus using numerical, quantitative format) in one trial and in conjunction with targeting (to Latinas versus no targeting) in another trial. The strength of the evidence in both studies was insufficient, however, rendering it impossible to draw definitive conclusions about the impact of any of these strategies (two trials; insufficient evidence).
- **Use of multiple communication strategies simultaneously:** In several cases, investigators used some combination of the four communication strategies when developing their interventions instead of comparing only a single strategy with another single strategy. Because comparisons were not one-to-one, it was more challenging to isolate the effects of each strategy.
- **Combining communication strategies with channel variation.** In one trial, investigators enhanced the communication strategy by *also* varying the communication channel for the intervention (i.e., using a lay health worker). While this tactic creates the potential for a more powerful effect, it is also complicates determining the effect of each strategy relative to the other.
- **Variation in use of strategies for patients versus clinicians.** None of the trials that met our review addressed using the four communication strategies with clinicians, therefore, we were unable to address KQ1b.

Table 13 documents the strength of evidence grading for each of the five comparisons and gives the overall SOE grade.

Table 13. Strength of evidence of communication strategies

Number of Studies; Subjects; Design/Quality	Risk of Bias	Consistency	Directness	Precision	Comparisons and Results Strength of Evidence
Framing (gain, loss) vs. Narratives (yes/no) (i.e., anecdotal evidence vs. statistical evidence)					
1; 174[75] RCT/Fair	Moderate	Consistency unknown (single study)	Direct	Imprecise	Framing (gain, loss) vs. narrative (yes/no) (i.e., anecdotal evidence/statistical evidence). RESULTS: Among those exposed to narrative/anecdotal evidence, the loss-framed messages were more persuasive than gain-framed messages. Among those exposed to non-narrative/statistical evidence, the likelihood of getting a mammogram did not differ significantly between those who received a gain-framed or a loss-framed message. Insufficient

Table 13. Strength of evidence of communication strategies (continued)

Number of Studies; Subjects; Design/Quality	Risk of Bias	Consistency	Directness	Precision	Comparisons and Results Strength of Evidence
Framing (gain, loss) vs. targeting (Latina targeting vs. multicultural targeting)					
1; 752[76] fRCT/Fair	Moderate	Consistency unknown (single study)	Direct	Imprecise	Framing (gain, loss) vs. targeting (Latina targeted vs. multicultural non-targeted): RESULTS: The loss-framed, multicultural message was statistically significantly more persuasive in terms of mammography use than the other three kinds of messages at 6 months. However, differences were not significant after 12 months. Insufficient
Targeting vs. Tailoring					
1; 1,574[77] Randomized trial/Fair	Moderate	Inconsistent	Direct	Precise	Tailoring vs. targeting: RESULTS: Letters that were personally tailored using an individual's risk-factor information did not increase the likelihood of receiving breast and cervical cancer screening among disadvantaged women. Rather, women receiving a targeted form letter with only standardized paragraphs about risk factors (age, gender) were more likely than those receiving the tailored letter to get screened for breast and cervical cancer. Low
1; 357[80] RCT/Fair	Moderate	Inconsistent	Direct	Imprecise	Targeting vs. tailoring [vs. tailoring plus lay health worker]. RESULTS: No statistically significant differences emerged in terms of percentage of calories consumed from fat or dietary fiber consumption among those exposed to the targeted or tailored version of the interventions. Adding a lay health worker intervention to the tailored version of the intervention did not have a statistically significant impact. Low
Targeting and tailoring vs. targeting only					
1; 5,500[78] RCT/Good	Low	Consistency unknown ((single study)	Direct	Precise	Targeting and tailoring vs. targeting: RESULTS The targeted and tailored intervention message did not lead to higher mammography rates than the targeted only group. The intervention was no more effective than the control group. Low

Abbreviations: RCT = randomized controlled trial.

Detailed Synthesis

Table 14 describes the five trials and their results in detail. This information is presented in ways reflecting the conceptual framework and the orientation of analysis of the original investigators. In most cases, however, the review team had to calculate differences between groups (e.g., in mean values on a scale or percentages). Because numerous findings led to negative differences (because of the original choices about the directionality of comparison), the table indicates whether the difference was negative or positive and notes which group the

findings favored. By favored, we mean which study group had the better result, namely a higher screening rate or better eating habit.

Table 14. Studies of communication strategies

Strategy	Author, Year Design Setting Sample Size Followup Quality Grade	Intervention Groups (N)	Outcomes	Results	Differences in Comparisons of Groups
Framing (gain, loss) vs. narrative (yes/no) (anecdotal evidence/ statistical evidence)	Cox and Cox, 2001[75] RCT US, Midwestern metropolitan areas Overall N=174 Immediate post-test Moderate	G1: Control (57) G2: Gain frame and non-narrative/statistical evidence (29) G3: Loss frame and non-narrative/statistical evidence (29) G4: Gain frame and narrative/anecdotal evidence (29) G5: Loss frame and narrative/ anecdotal evidence (29)	Self-reported predicted likelihood of getting a mammogram after seeing advertisement, comparing "loss" to "gain" frame holding type of evidence (narrative/not narrative) constant; and looking at their interactive effect	Mean likelihood (7-point Likert scale where a higher number means greater likelihood): G2: 5.48 G3: 4.37 G4: 4.07 G5: 5.54	G3 vs. G2: -1.11[a] Difference not significant, p not reported (favoring G2) G5 vs. G4: 1.47[a] (p<0.01) (favoring G5) Using 2x2 ANOVAs: Framing x evidence interaction: F(1, 103)=10.87, p=0.001 <u>Narrative/Anecdotal evidence:</u> Loss frame statistically more effective than gain frame: F(1, 103)=7.57, p<0.01 <u>Non-narrative/Statistical evidence</u>: No statistically significant difference between loss and gain frame : F(1, 10-3)=3.77, p=.06
Framing (gain, loss) vs. targeting (Latina targeting vs. multicultural non-targeted)	Schneider et al., 2001[76] RCT (factorial design) US, community health clinics, and public housing N=752 at 6 months N=444 at 12 months 6 and 12 months	G1: Gain frame and non-targeted (multicultural) G2: Loss frame and non-targeted (multicultural) G3: Gain frame and targeted (to Latinas) G4: Loss frame and targeted (to Latinas) Group sizes not reported	Self-reported likelihood of getting a mammogram in the past 12 months within 6 months after seeing video, comparing loss to gain frames holding type of targeting constant; and looking at their interactive effect	Percentage Overall: 41% G1: 36% G2: 50% G3: 41% G4: 36%	G2 vs. G1: 14%[a]=odds ratio of 1.81 (p<0.01) (favoring G2) G4 vs. G3: -5%[a]= odds ratio of 1.22 (p>0.10) (not significant) Using hierarchical logistic regression, controlling for past year's use: Framing x targeting interaction: Chi-square=5.15, p<0.05

Table 14. Studies of communication strategies (continued)

Strategy	Author, Year Design Setting Sample Size Followup Quality Grade	Intervention Groups (N)	Outcomes	Results	Differences in Comparisons of Groups
Framing (gain, loss) vs. targeting (Latina targeting vs. multicultural non-targeted) (continued)	Fair				Odds ratios (CIs): Past year's Mammography use: 1.44 (0.98 to 2.11) Framing (loss): 1.27 (0.78-2.08) Targeting: 1.20 (0.72 to 1.99) Frame x Target: 2.27 (1.12 to 4.63)
			Self-reported likelihood of getting a mammogram in the past 12 months within 12 months after seeing video, comparing "loss" to "gain" frames holding type of targeting constant; and looking at their interactive effect	Percentage Overall: 57% G1: 55% G2: 61% G3: 57% G4: 54%	G2 vs. G1: 6% [a] Not significant, p value not reported (favoring G2) G4 vs. G3: -3% [a] Not significant, p value not reported (favoring G3) Using hierarchical logistic regression, controlling for past year's use: Framing x targeting interaction: Chi-square=1.65, not significant Odds ratios (CIs): Past year's Mammography use: 2.93 (2.05 to 4.18), p<0.01 Framing (loss): 1.18 (0.74 to 1.89) Targeting: 1.05 (0.65 to 1.70) Frame x Target: 1.56 (0.79 to 3.08)

Table 14. Studies of communication strategies (continued)

Strategy	Author, Year Design Setting Sample Size Followup Quality Grade	Intervention Groups (N)	Outcomes	Results	Differences in Comparisons of Groups
Targeting versus tailoring	Jibaja-Weiss et al., 2003[77] RCT US Community health clinics in Houston, Texas, that provide care to underserved and low-income neighborhoods N=1574 12 months Fair	G1: No intervention control (499 for cervical, 239 for breast) G2: Personalized form letters targeted to women age 40 and older (460 for cervical, 239 for breast) G3: Personalized tailored letter (524 for cervical, 261 for breast)	Self-reported likelihood of scheduling a cervical cancer screening test appointment within 12 months after exposure to the letter	Percentage G1: 44.7% G2: 53.3% G3: 39.7%	G3 vs. G2: -13.6 [a] (favoring G2) Overall Chi-square test (comparing all 3 groups) p<0.001
			Self-reported likelihood or receiving a cervical cancer screening test within 12 months after exposure to the letter	G1: 39.9% G2: 43.9% G3: 23.7%	G3 vs.G2: -20.2% [a] (favoring G2) Overall Chi-square test (comparing all 3 groups) p<0.001
			Self-reported likelihood of scheduling breast cancer screening test appointment within 12 months after exposure to the letter	G1: 53.3% G2: 65.7% G3: 50.2%	G3 vs. G2: -15.5% [a] (favoring G2) Overall Chi-square test (comparing all 3 groups) p= 0.001
			Self-reported likelihood of receiving a breast cancer screening test within 12 months after exposure to the letter	G1: 20.7% G2: 30.5% G3: 13.0%	G3 vs. G2: -17.5% [a] (favoring G2) Overall Chi-square test (comparing all 3 groups) p<0.001

Table 14. Studies of communication strategies (continued)

Strategy	Author, Year Design Setting Sample Size Followup Quality Grade	Intervention Groups (N)	Outcomes	Results	Differences in Comparisons of Groups
Targeting vs. vs. tailoring vs. tailoring plus lay health worker	Elder et al., 2005[80] RCT US, San Diego County, with dominant Latino populations Overall N=357 12 weeks Fair	G1: Control materials targeted to a Latino population ("off the shelf" materials covering same modules and content as lay health workers) (119) G2: Tailored print materials (118) G3: Tailored print materials plus lay health worker (120)	Calories from fat (percentage)	Percentage at baseline minus percentage at 12 weeks G1: 31.5-30.0=1.5 G2: 31.0-30.4=0.6 G3: 31.5-29.3=2.2	Difference of differences between G2 vs. G1: -0.9 [a] (favoring G1=fewer calories from fat) Differences among the 3 groups at 12 weeks controlling for baseline level not significant F=0.81, p=0.45,
			Total dietary fiber (grams)	Adjusted mean, in grams at baseline minus grams at 12 weeks G1: 16.5-15.6=0.9 G2: 17.2-17.2=0.0 G3: 17.2-16.1=1.1	Difference of the differences between G2 vs.G1: -0.09 [a] (favoring G1 = more grams of fiber) Differences among the 3 groups at 12 weeks controlling for baseline level not significant F=1.61, p=0.20, not significant
Targeting and tailoring vs. targeting only	Vernon et al., 2008[78] RCT US , National Registry of Women Veterans Overall N=5500 12 and 24 months Good	G1: No intervention control (1,840 for 12 months, 754 for 24 months) G2: Targeted (1,857 for 12 months, 825 for 24 months) G3: Targeted and tailored (1,803 for 12 months, 781 for 24 months)	Self-reported likelihood of getting a breast cancer screening within 12 months after exposure to the letter	Crude incidence (percentage) G1: 44.7% G2: 46.9% G3: 46.0%	ITT difference G3 vs. G2: -0.9% a (not significant) Chi-square=1.70, 2 d.f. p=0.427 Using Cox proportional hazard rate ratio (CI) using ITT: Differences not significant. G1: 1.00 G2: 1.07 (0.97 to 1.18) G3: 1.05 (0.95 to 1.15)
			Self-reported likelihood of getting a breast cancer screening within 24 months after exposure to the letter	Crude incidence (%): G1: 22.0% G2: 24.8% G3: 24.8%	ITT difference G3 vs. G2: 0.0% a Chi-square=5.17 2 d.f. p=0.075 (G2 and G3 are equal) Cox proportional hazard rate ratio (CI) using modified ITT: Differences not significant. G1: 1.00 G2: 0.99 (0.86 to 1.13) G3: 1.05 (0.91 to 1.20)

[a] Calculated by reviewers.

Abbreviations: CI = confidence interval; d.f. = degrees of freedom; F = F-test; G = group; ITT = intention to treat; N = number; RCT = randomized controlled trial; US = United States; vs. = versus

Framing (Gain, Loss) Versus Narratives (Yes/No) (Anecdotal/ Statistical Evidence)

A small (N = 174) trial examined the effect of experimental advertisements that differed in terms of how consequences of getting screened for breast cancer (with mammography) were framed (gain or loss) and how the evidence was presented (narrative/anecdotal vs. non-narrative/statistical) (Table 14).[75] In this trial, investigators randomly assigned 116 women to one of four groups with different message combinations: (1) control; (2) gain frame and non-narrative/statistical; (3) loss frame and non-narrative/statistical; (4) gain frame and narrative/anecdotal, and (5) loss frame and narrative/anecdotal. Gain-framed messages focused on the potential for screening to save lives ("…they are less likely to die of breast cancer"); loss-framed messages focused on the possibility of death from not being screened ("…they are more likely to die of breast cancer"). The narrative/anecdotal approach involved personal narrative stories ("Doctors were able to detect her breast cancer at an early, treatable stage, and now Sara can look forward to a long life, watching her grandson, Jeffrey, grow up"); non-narrative/statistical approach had a numerical emphasis ("Doctors are able to detect their tumors at an early, treatable stage, and they [women] are 30% less likely to die of breast cancer").

The effect of the different approaches varied. Women who received the narrative/anecdotal and loss-framed message (Group 5) reported the highest mean likelihood of getting a mammogram (Group 5; 5.54 on a 7-point Likert scale). Likelihood values dropped off as follows: non-narrative/statistical and gain frame (Group 2; 5.48); non-narrative/statistical and loss frame (Group 3; 4.37); and finally narrative/anecdotal and gain frame (Group 4; 4.07). Framing and use of narratives had an interactive effect on subjects' predictions of their own mammography behavior. The effects of framing were moderated by how the evidence was presented; specifically, among those exposed to narrative/anecdotal evidence, the loss-framed messages were more persuasive. By contrast, among those receiving non-narrative/statistical evidence rather than narrative/anecdotal information, the likelihood of getting a mammogram did not differ significantly between those who received gain-framed and a loss-framed messages ($p=0.06$).

Based on this single study, we graded the overall strength of evidence as insufficient because of the small sample sizes in each intervention group (n = 29), the use of a convenience sample (reflected in the risk of bias assessment of moderate), and the imprecision of the results that were evaluated only immediately after exposure to the intervention (precluding any conclusions).

Framing (Gain, Loss) Versus Targeting (Latina Targeted Versus Multicultural)

One trial used a 2 x 2 factorial design to examine the effect of message framing (gain vs. loss) and targeting on the basis of ethnicity of the women receiving the information; the two options were targeting Latinas only or taking a multicultural orientation (women of various ethnic backgrounds).[76] The purpose of the messages was to motivate breast cancer screening in low-income women who are medically underserved. The investigators hypothesized that targeting would enhance attention to the message, especially for the loss frame. They showed women older than 40 one of four videos with four different message strategies: (1) gain frame and non-targeted/multicultural, (2) loss frame and non-targeted/multicultural, (3) gain frame and targeted toward Latinas, and (4) loss frame and targeted toward Latinas. Participants self-reported information about mammography use 6 and 12 months after exposure to the videos. The

investigators contacted study participants by either telephone or mail (stamped, preaddressed envelope).

On average, 41 percent of participants reported having a mammogram *within the past 12 months* 6 months after exposure to the intervention. Those who received the loss frame and non-targeted (multicultural) video reported the highest percentage of mammograms (Group 2; 50%), followed by those exposed to the gain frame and Latina targeted video (Group 3; 41%). Among women receiving the gain frame and non-targeted (multicultural) video and those receiving shown the loss frame and Latina targeted video (Groups 1 and 4), 37% of participants had a mammogram within the past 12 months 6 months after exposure to the intervention. Using hierarchical logistic regression, controlling for the past year's screening usage, framing and targeting had a significant interactive effect on the probability of getting a mammogram. "Within 6 months after participation, loss framed videos persuaded more participants to obtain mammograms than the gain framed videos, but only among those viewing the non-targeted multicultural context. Unexpectedly, the loss framed, non-targeted/multicultural message was more persuasive in terms of mammography use than the other three kinds of messages."[76, p.260]

On average, 57 percent of participants reported having a mammogram *within the past 12 months* 12 months after exposure to the intervention. The pattern of differences among study groups was similar at 6 and 12 months. Within 12 months after participation, and controlling for the past year's screening usage, the framing/targeting interaction in the logistic regression model was not statistically significant. Based on this single trial of relatively large size, moderate risk of bias and imprecise estimates, we graded the strength of the evidence as insufficient.

Targeting Versus Tailoring

One trial evaluated the effectiveness of two types of letters to encourage low-income women served by county public health clinics to adhere to cancer screening recommendations.[77] One group of women (Group 2, denoted PF) received a form letter targeted to women age 40 and older; another group of women (Group 3, denoted PT) received a personally tailored letter (based on information from their medical record about their personal risk of cancer) (e.g., "Mrs. Smith, you may be at risk of breast cancer because…); and a third group (Group 1) served as a control. Both interventions sought to prompt women to get screened for breast and cervical cancer.

Unexpectedly, women who received the personally tailored intervention (Group 3, PT) were *less* likely to schedule (13.6 percentage points less) or obtain (20.2 percentage points less) cervical screening within the first 12 months after receiving the letter than women who received targeted form letter (Group 2, PF). Percentages for control group members fell between the targeted and tailored groups. The differences among the three study groups for breast cancer screening were generally comparable to those for cervical cancer screening.

Another trial examined innovative approaches to changing lifestyle behaviors to reduce dietary fat and to increase fiber among Latinas and their families in two counties in Southern California.[80] One intervention group (the "tailored print" material option) received 12 weekly newsletters tailored using baseline survey data provided by the participants. The newsletters contained activity inserts and a story (novella) about a woman gaining control of her personal life. This group also received supporting materials (e.g., recipes and magnets to prompt behavior change). All materials were delivered to women in their homes via the US postal service. The other intervention group (the "lay health worker/promotora tailored print material" option) received weekly home visits or telephone calls from promotoras (lay health workers in their community) over a 12-week period plus the 12 newsletters and activity inserts also delivered by

mail. Finally, a control group received "off the shelf" materials by mail covering the same modules and content as the intervention groups; these women received no tailoring or personal interaction, but the materials were targeted to Latinas.

Using analysis of covariance, controlling for baseline levels, the investigators found no statistically significant differences in terms of a percentage point decrease in calories consumed from fat over a 12 week period among the three groups. No statistically significant differences emerged in terms of the difference in percentage point decrease in dietary fiber consumption among the three groups. Based on these two trials with moderate risk of bias and varying levels of precision and inconsistent findings, we graded the strength of the evidence as low.

Targeting and Tailoring Versus Targeting Only

One large (N=5,500) trial developed and evaluated interventions to promote breast cancer screening in a nationally representative sample of U.S. women veterans.[78] The trial included women ages 52 and older who were no longer on active duty. One group (targeted only) received a less personalized, targeted intervention; it included a letter for women to use to discuss mammography with her health care provider, educational booklets, and a pamphlet about mammography screening services available through the Veterans Health Administration (VA). Another group (targeted and tailored) received the same materials, but their tailored letter consisted of individualized messages that addressed each participant's responses to questions on a baseline survey reflecting her attitude and opinions about screening, including her stage of readiness to be screened. A control group received no intervention. All groups completed the baseline survey.

The more intensive intervention (tailored and targeted) was no more effective than the less intensive intervention (targeted only) in terms of screening rates at 12 months (46.0% vs. 46.9% screened, respectively) and at 24 months after the intervention (24.8% vs. 24.8% screened, respectively). Further, differences between the both intervention groups and the control group at both time periods were not statistically significant based on multivariate analyses (namely, Cox proportional hazard modeling and intention to treat [ITT] analyses). Modified ITT was used to evaluate screening at 24 months because it was conditional on having a mammogram at 12 months. This single trial had low risk of bias and precise estimates given adequate power and thorough analyses, and we graded the strength of the evidence as low.

Results: Key Question 2: Dissemination Strategies to Clinicians and Patients

Introduction

The analytic framework presented in the Introduction guided the KQ2 literature search for Key Question (KQ) 2 abstraction and analysis. KQ2 concerns dissemination to clinicians or patients. As discussed in the Methods section, we included dissemination strategies that met two main criteria. First, they attempted to increase reach, ability, or motivation or used multi-component approaches addressing one or more of these strategies. We categorized specific dissemination tactics as comprising reach, ability, motivation or multi-component approaches as described previously. Second, they attempted to address health-related decisions and behaviors, clinical outcomes, or knowledge. We included only randomized controlled trials (RCTs) or cluster RCTs (cRCTs)—hereafter referred to as trials.

This section presents the results that compare the dissemination strategies noted above. Some trials compared strategies directly with each other (e.g., ability strategies vs. motivation strategies) and can be regarded as head-to-head trials for comparative effectiveness analyses. Some trials compared strategies to a usual care or no-treatment control group, but we included them in our analysis if they had at least two trial arms that addressed our inclusion criteria and if we believed that we might glean information about the relative effectiveness of one strategy versus another. In many cases where there was not a direct comparison significant tests or confidence intervals were likely also not reported, and we note this in the summary tables.

We divided the trials by dissemination strategies and by outcomes for clinicians and patients. The included studies were very heterogeneous with regard to the specific health-related decisions and behaviors, clinical outcomes or knowledge studied. They were also very heterogeneous with regard to specific dissemination strategies reflecting the reach, ability, motivation, and multi-component approaches. Because of this heterogeneity, we chose to present summary tables in this section that describe trial results in words rather than numerically. Information reflects the comparisons as presented by the original investigators, which in some cases do not reflect precisely the constructs or directionality of our conceptual framework and definitions of strategies. In addition, in grading the strength of evidence (SOE), we focused on the primary outcomes as defined by the authors of the papers reporting on the trials. Detailed evidence tables presenting numerical findings for abstracted trials appear in Appendix E.

Figure 1 in the previous section presents the flow diagram that accounts for all titles/abstracts and full-text articles reviewed and the final set of articles for all three KQs. In all, 153 articles were relevant to KQ2. After full-text review, we retained 40 articles that met inclusion criteria; they report on 37 unique trials (including one trial that reported on different outcomes in two articles) about dissemination strategies.

Description of Included Studies

Two team members independently assessed study quality for 153 articles related to KQ2. When there were disagreements they were adjudicated by 2-3 team members. We used 37 trials in the KQ2 analysis. Of these, we assessed 15 as good quality (i.e., low risk of bias)[81-96] and 22 as fair quality (i.e., moderate risk of bias).[80,97-117] Those assessed as poor quality (i.e., high risk of bias) were not abstracted or used for analysis.

These 37 trials reported a wide variety of primary and secondary outcomes that spanned a range of health-related or clinical problems; these are documented in the outcome column in Tables 14-30. Even though some trials examined similar outcomes, e.g., guideline-concordant care, they may have used different dissemination strategies or focused on different problems or clinical issues. Because of this heterogeneity, we were not able to combine trials to conduct a quantitative analysis of the effect of dissemination to clinicians or patients.

We graded the SOE for trials at the level of dissemination strategy (increasing reach, ability, motivation, or multi-component) and outcome category (health-related decision or behavior, clinical outcome or knowledge). In some cases, only one trial was relevant to the analysis (i.e., a single study for which SOE could be graded); in other cases we had as many as six trials on a strategy-outcome topic, but they were different enough to preclude any meta-analysis.

The trials were conducted in the United States, Canada, England, Germany, Finland, the Netherlands, Scotland, and Spain. Sample sizes ranged from 114 participants to 3,293 participants. For the cRCTs, cluster sizes ranged from 9 to 249.

In the tables below (Tables 14-30), we summarize the intervention comparisons that the authors had used along with our classification of specific strategies attempting to affect reach, ability, motivation or a combination of these aims. The summary tables show the following: author, year of publication, study design (RCT or cRCT), number of study participants (N) and/or number of clusters (N_c) randomized (if reported), study duration, quality rating for the trial; setting and sample characteristics; study intervention groups (e.g., G1,…GX); and the outcomes as defined by the authors; and a summary of the trial's results.

We described summary results in several ways. If one group (e.g., G1) had a stronger effect than another group (e.g., G2) we denoted this as G1>G2. If significance tests or confidence intervals were reported for the direct comparison, and they indicated statistical significance, we described these in the table as *significant differences*. In some cases there were no significance tests or confidence intervals reported for the comparisons of interest; instead, the active arms were compared to a control group and we note this in the tables as no significance tests or confidence intervals reported for the comparison of interest.

If groups did not differ we noted this in summary tables as G1=G2. Here again, if statistical significance tests or confidence intervals were reported for the direct comparison we described these as *no significant differences*. As previously mentioned, in some cases there were no significance tests or confidence intervals reported. In these cases we noted this in the tables no significance tests or confidence intervals reported for the comparison of interest.

In the next three sections we review the evidence for dissemination strategies focused on clinicians and patients. In the first section we focus on clinicians, in the second section on patients, and in the third section on studies that targeted both clinicians *and* patients. In each section summarize key points, present the strength of evidence table for each comparison category and summarize the evidence in tables and the narrative. The summary tables describe the study and key findings. The narrative describes additional details, including the specific dissemination tactics a trial used.

Key Question 2: Disseminating Evidence to Clinicians

Key Points: Clinician Trials

- Ability strategies not more effective than reach strategies related to clinician behavior (4 trails, low SOE).

- One study found that motivation strategies may be more effective than reach strategies for influencing clinicians' behaviors (1 trial; insufficient evidence).
- Multi-component strategies that address a combination of reach, ability, or motivation appear to be more effective than one strategy alone for affecting clinicians' behaviors, particularly guideline adherence (6 trials; moderate SOE) and for clinical outcomes although many comparisons examining clinical outcomes were not significant (4 trials, low SOE.)
- The strength of evidence is insufficient for most comparisons related to clinical outcomes and knowledge for clinicians because we had only single trials in each case.

Detailed Synthesis: Dissemination Strategies for Clinicians and Health-related Decisions or Behaviors

Table 15 documents the strength of evidence grading for each of the dissemination comparisons focused on clinicians for health-related decisions and behaviors and gives the overall SOE grade.

Table 15. Strength of evidence: dissemination to clinicians for health-related decisions and behavior outcomes

Strategy	Number of Trials; Subjects; Design/Quality Rating	Risk of Bias	Consistency	Directness	Precision	Results and Strength of Evidence
Reach strategies vs. ability strategies	4; N_c = 199 N=213 3 cRCT/Good[81,88,94] 1 RCT/Fair[115]	High	Consistent	Direct	Imprecise	Ability strategies not better than reach strategies. All studies found no significant differences between groups receiving these two approaches. Low for no differences
Reach strategies vs. motivation strategies	1; N =363; RCT/Fair[105]	Moderate	Unknown (single study)	Direct	Imprecise	Motivation strategy significantly better than reach strategy. Insufficient
Reach strategies vs. multi-component strategies	6; N_c= 237; N=3,860 3 cRCT/Good[83,93,95] 3 RCT/Fair[97,109,117]	Moderate	Consistent	Direct	Precise	Multi-component strategies better than reach strategies. Almost all studies using multi-component strategies found significant differences between groups receiving these two approaches. Moderate for multi-component strategies
Ability strategies vs. multi-component strategies	3; N_c = 382 N=2624 2 cRCT/Fair[106,112] 1 RCT/Fair[114]	Moderate	Consistent	Direct	Precise	Results appear to favor multi-component strategy over ability strategy. However, no confidence intervals or statistical test reported for comparisons between arms for one of the trials. Low

Table 15. Strength of evidence: dissemination to clinicians for health-related decisions and behavior outcomes (continued)

Strategy	Number of Trials; Subjects; Design/Quality Rating	Risk of Bias	Consistency	Directness	Precision	Results and Strength of Evidence
Motivation strategies vs. multi-component strategies	1; N_c = 42 cRCT/Good[96]	Low	Unknown (single study)	Direct	Precise	Multi-component strategy significantly better than motivation strategy. Insufficient
Reach strategies vs. ability strategies vs. multi-component strategies	1; N_c = 61 cRCT/Fair[116]	Moderate	Unknown (single study)	Direct	Imprecise	Multi-component strategy not better than either reach or ability strategies Insufficient

Abbreviations: cRCT = clustered randomized controlled trial; N = number; N_c = number of clusters; RCT = randomized controlled trial

Table 15 describes the clinician-focused trials and summarizes their results for decision or behavioral outcomes. The text accompanying the table describes trials that compare (a) reach and ability strategies, (b) reach and motivation strategies, (c) reach versus multi-component strategies, (d) ability versus multi-component strategies, (e) motivation versus multi-component strategies, and (f) ability, or motivation and multi-component strategies.

Strategies to Increase Reach versus Strategies to Increase Ability

Four trials tested dissemination strategies to affect clinician health-related decisions and behaviors by increasing reach or increasing ability (Table 16). Three were cRCTs,[81,88,94] and one was a RCT.[115] Reach strategies included delivering guidelines by mail or computer; increasing ability strategies included computer assisted learning, textbooks, and individual or group academic detailing. The follow-up periods ranged from 1 month to 2 years after the intervention. All trials examined guideline-concordant care, but for different behaviors and conditions. Across these three trials, intervention groups did not differ significantly for any of the primary outcomes reported.

Strategies to Increase Reach Versus Strategies to Increase Motivation

One trial tested dissemination strategies to affect clinician behavior by increasing reach versus increasing motivation (Table 16).[105] This trial compared conventional guideline delivery with a scheme that contained vignettes involving patient-specific ratings from an expert panel as part of the guideline message. For example "the expert panels recommend exercise testing (rating 7)," where the rating corresponded to the appropriateness of that test for a specific patient portrayed in a vignette (e.g., patient specific ratings). The outcome measure was the agreement between the physician's test ordering and recommendations of two independent expert panels. Immediate post-test measures after the intervention showed providing this vignette-driven, patient-specific information about the appropriateness of applying a guideline was significantly more effective than simply providing traditional guidelines.

Table 16. Summary of trials examining dissemination to clinicians for health-related decisions and behavior outcomes

Author, Year, Design Sample Size Duration Quality Grade	Setting Sample Characteristics	Intervention Groups	Outcome	Results
Trials that compared reach strategies with ability strategies				
Bahrami et al., 2004[81] cRCT N_c = 51 Two 4-month periods pre and 2 4-month periods post intervention Good	Clinical Practices associated with the Scottish Dental Practice Board.	G1: Mailed guideline (increase reach) G2: Guideline + audit and feedback (AF) (not abstracted) G3: Computer-assisted learning (CAL) (increase ability) G4: CAL + AF (not abstracted)	Guideline compliance (proportion of patients whose treatment complied with the guideline) (P)	Confidence intervals or p values not reported for this comparison
Jousimaa et al., 2002[88] cRCT N_c = 139 1-month post-intervention Good	Clinical Recently qualified physicians who would work in a Finnish health center for at least 2 months	G1: Computerized version of guidelines (increase ability) G2: Textbook-based version of guidelines (increase reach)	Physicians' compliance with guideline recommendations about laboratory, radiologic, physical, and other examinations; procedures; nonpharmacologic and pharmacologic treatments; physiotherapy; and referrals (P)	G1 = G2 No significant difference between groups
Simon et al., 2005[94] cRCT N_c = 9 N (clinicians) = 781 N (patients) = 9,820 Baseline, 1-year follow-up, 2-year follow-up Good	Clinical All patients with hypertension receiving primary care at 1 of 9 study sites; all clinicians providing primary care for adults at the 9 study sites	G1: Mailed educational materials (increase reach) G2: Individual academic detailing (increase ability) G3: Group academic detailing (increase ability)	Change in guideline adherence (proportion of patients with new hypertension receiving a diuretic or beta blocker) (P)	G1 = G2 G1=G3 No significant difference among groups Confidence intervals or p values not reported for G2 vs. G3 comparison

Table 16. Summary of trials examining dissemination to clinicians for health-related decisions and behavior outcomes (continued)

Author, Year, Design Sample Size Duration Quality Grade	Setting Sample Characteristics	Intervention Groups	Outcome	Results
Sullivan et al., 2010[115] RCT N = 213 Baseline, immediate posttest, 60-day posttest Fair	Clinical Residents in internal medicine	G1: VA guidelines for pain management (increase reach) G2: COPE: web-based education program for pain management (increase ability)	Use of four core management strategies for pain management (2nd)	G1=G2 No significant difference between groups
Trials that compared reach strategies with motivation strategies				
Junghans et al., 2007[105] RCT N = 363 Immediate posttest Fair	Clinical Members of the British Cardiac Society or general practioners in the primary care trusts that referred patients to 9 cardiothoracic centers in England and Scotland	G1: Conventional guideline (increase reach) G2: Ratings about specific patients in vignettes (increase motivation)	Physician testing behavior (agreement of physicians' recommendations with those made by 2 independent expert panels) (P)	G2 > G1 Patient-specific ratings endorsed by experts significantly more effective than conventional guidelines.
Trials that compared reach strategies with multi-component strategies				
Banait et al., 2003[117] RCT N = 114 7 months pre and post intervention Fair	Clinical All general practices in Greater Manchester, England	G1: Mailed guidelines (increase reach) G2: Educational outreach (Multi-component)	Uptake of dyspepsia management guidelines, including: • Appropriateness of referrals for open access endoscopy (P) • Findings at open access endoscopy (P)	G2 > G1 Educational outreach significantly more effective than postal dissemination for appropriateness of referrals for open access endoscopy G2=G1 for findings at open access endoscopy

Table 16. Summary of trials examining dissemination to clinicians for health-related decisions and behavior outcomes (continued)

Author, Year, Design Sample Size Duration Quality Grade	Setting Sample Characteristics	Intervention Groups	Outcome	Results
Beaulieu et al., 2004[97] RCT N = 3293 6 month follow-up Fair	Clinical Quebec physicians who had to be a primary prescribing physician (responsible for more than half of all anti-anginal prescriptions) for at least one patient (moe than 65 years of age)	G1: Control (not abstracted) G2: Guideline (increase reach) G3: Guideline + reminder notice and stickers for patients' charts (multi-component)	Prescribing practices in line with guildelines (prescription of 3 cardiovascular medications) (P)	Confidence intervals or p values not reported for comparisons
Bekkering et al., 2005[83] cRCT N_c = 68 N (physiotherapist) = 113 N (patient) =511 Baseline unspecified follow-up Good	Clinical Physcial therapy practices that were members of the Royal Dutch Society for Physical Therapy in the Netherlands whose physiotherapists expected to treat at least 5 patients with low back pain during the enrolment period.	G1: Mailed guideline (improve reach) G2: Multi-faceted (multi-component)	Guideline adherence: (P) • Proportion of patients for whom each and all 4 guidelines were fullfilled	G2 > G1 Arm with multifaceted education, role playing and reminders (multi-component) significantly more effective than arm with paper-based mailed guideline dissemination
Murtaugh et al., 2005[109] RCT N = 354 Within 45 days after initial home health RN assessment Fair	Clinical English- or Spanish-speaking patients, ages18 or older, with a primary diagnosis of heart failure	G1: Usual care (not abstracted) G2: Basic intervention e-mail reminder (increase reach) G3: Augmented intervention of e-mail reminder + package of supporting materials (multi-component)	Practice of evidence-based care: (P) • Recording key assessment items and instructions to patients • Instructing patients with key educational elements related to symptoms and side effects • Assessment of patient diet and medication side effects	Confidence intervals or p values not reported for comparison

Table 16. Summary of trials examining dissemination to clinicians for health-related decisions and behavior outcomes (continued)

Author, Year, Design Sample Size Duration Quality Grade	Setting Sample Characteristics	Intervention Groups	Outcome	Results
Rebbeck et al. 2006[93] cRCT N_c = 27 N = 99 Baseline, 1.5 months, 3 months, 6 months, and 12 months Good	Clinical Physiotherapy clinics in Australia Clinic: Clinics in 2 states in Australia that had seen at least 5 whiplash cases in the previous year Patient: Patients, 18 years and older, involved in a motor vehicle accident within the previous 6 weeks who presented to the clinic with acute whiplash	G1: Dissemination of guidelines by mail (increase reach) G2: Implementation group (multi-component)	Physiotherapist prescribing correct treatments (medication, exercise, and function) (2nd) Advising to act as usual (2nd)	G1=G2 Physiotherapist prescribing correct treatments (medication, exercise, and function) no significant differences between groups G2>G1 Multi-component arm more effective than reach strategy for advising to act as usual
Wolters et al., 2005[95] cRCT N_c =142 Baseline, up to 1 year post intervention Good	Clinical Patients: Men with lower urinary tract symptoms older than 50 years visiting partipating clinics in the Netherlands	G1: Control mailed guidelines (increase reach) G2: Intervention involving package for learning, supporting materials, decision tree, and information leaflets for patients (multi-component)	General practice management of men with lower urinary tract symptoms: • Number of PSA requests (P) • Medication prescribed (P) • Referral rate to a urologist (P) • Provision of patient education materials (2nd) • Discussion of the PSA evidence (2nd)	G2 > G1 Multi-component strategy significantly more effective than paper-based mailed guidelines dissemination for referral rates to a urologist and provision of patient education materials G1=G2 No difference between arms for number of PSA requests, medication prescribed, and discussion of the PSA evidence

Table 16. Summary of trials examining dissemination to clinicians for health-related decisions and behavior outcomes (continued)

Author, Year, Design Sample Size Duration Quality Grade	Setting Sample Characteristics	Intervention Groups	Outcome	Results
Trials that compared ability strategies with multi-component strategies				
Laprise et al., 2009[106] cRCT N_c = 133 N = 2,344 Baseline; 6-month follow-up Fair	Community-based settings All practicing GPs in 5 regions of Quebec, Canada, who had to see at least 25 patients 55 years of age or older in any 2- week period.	G1: CME (increase ability) G2: CME + practice enablers and reinforcers (multi-component)	Guideline adherence: • Proportion of patients receiving at least one prescription for managing cardiovascular disease (P) • Prescription of anti-platelets, angiotensine converting ensyme inhibitor, lipid lowering agents, beta blockers for undermanaged patients (2nd)	G2 > G1 CME plus practice enablers and reinforcers significantly more effective than only CME for proportion of patients receiving at least one prescription, prescription, and for angiotensine converting enzyme but not for anti-platelets, beta blockers, or lipid lower agents
Rahme et al., 2005[112] cRCT N_c = 249 135 to 1 day pre-intervention; 1 to 136 days post intervention Fair	Clinical GP: All GPs registered in 1 of 8 selected towns in Canada Patients: All patients 65 years or older who who had at least 1 diagnosis for osteoarthritis and filled a prescription for an NSAID, COX-2 inhibitor, or acetaminophen written by a GP in the study	G1: No treatment control (not abstracted) G2: Decision tree (increase ability) G3: Workshop (increase ability) G4: Workshop + decision tree (multi-component)	Dispensed prescriptions (prescription adequacy) (P)	Confidence intervals or p values not reported for comparison
Soler et al., 2010[114] RCT N = 2624 45 days post training and continued for 3 months Fair	Clinical GP: Physicians working at 1 of 40 general practices in Spain	G1: Control (not abstracted) G2: Training session on the SEPAR guidelines (increase ability) G3: G2 + portable device for spirometry (multi-component)	Management of COPD per guidelines: (P) • Improved diagnosis • Severity classification • Treatment regimen	Confidence intervals or p values not reported for comparison.

Table 16. Summary of trials examining dissemination to clinicians for health-related decisions and behavior outcomes (continued)

Author, Year, Design Sample Size Duration Quality Grade	Setting Sample Characteristics	Intervention Groups	Outcome	Results
Trials that compared motivation strategies with multi-component strategies				
Wright et al., 2008[96] cRCT N_c = 42 360 days before intervention, 360 days after intervention Good	Clinical Hospitals in Ontario, Canada, that identified a local opinion leader in colon cancer	G1: Standardized lecture by expert opinion leader (increase motivation) G2: Standardized lecture by expert opinion leader + academic detailing and a toolkit (multi-component)	Guideline adherence: (P) • Mean number of lymph nodes assessed in patients with stage II colon cancer • Number of lymph nodes removed	Inconsistent results. G2 > G1 : Lymph node assessment G2 = G1 : Lymph node removal Adding academic detailing and a toolkit to a standardized lecture by an expert opinion leader significantly more effective than standard lecture by expert opinion leader for lymph node assessment. No significant differences between groups for lymph node removal.
Studies that compared reach strategies, ability strategies, and multi-component strategies				
Watson et al., 2002[116] cRCT N_c = 61 Baseline, immediate posttest Fair	Community-based settings NR	G1: Guideline materials by postal mail (increase reach) G2: Education outreach (EO) session and guidelines (increase ability) G3: Continuing professional education (CPE) session and guidelines (increase ability) G4: Guidelines + EO and CPE (multi-complonent)	Guideline adherence: (P) • Proportion of visits resulting in an appropriate sale of an anti-fungal product	G1 = G2 = G3 = G4 No significant differences among groups

Note: In most cases, the labels of the intervention groups represent the labels that study authors had originally used, except if they were merely labeled control or intervention. Information in parentheses represents the strategy that the authors of this review believed best captured the intervention components. When authors used nondescriptive or vague labels for the intervention groups (e.g., Group 1, Intervention A) we used a summary of their description of the group as a label. For a full description of the studies intervention components refer to Appendix E.

Abbreviations: 2nd = secondary outcome; CME =, continuing medical education; COX-2 = Cyclooxygenase-2 inhibitor; GP = general practitioner; NR = not reported; NSAID = nonsteroidal anti-imflammatory drug; P = primary outcome; PSA = prostate-specific antigen; RN = registered nurse; + = plus.

Strategies to Increase Reach Versus Multi-Component Strategies

Six trials tested dissemination strategies to affect clinician behavior by increasing reach against multi-component strategy (Table 16).[83,93,95,97,109,117] Three trials were cRCTs. Each trial compared multi-component approaches that involved some type of educational component and written or in-person reminders or decisions support materials with reach strategies that entailed postal or emailed guidelines. One study used an opinion leader delivering guidelines as the reach strategy.[93] The follow-up periods ranged from 45 days to 1 year after the intervention. Across these six trials, five of the six found that multi-component strategies were more effective at changing outcomes compared to reach strategies. In the remaining trail[97] statistical significance tests were not reported for this comparison.

Strategies to Increase Ability Versus Multi-Component Strategies

Three trials tested dissemination strategies to affect clinician behavior by increasing ability against a multi-component approach (Table 16).[106,112] One was a cRCT with a follow-up period of up to 136 days after the intervention assessing whether a workshop that provided education about evidence-based medicine and decision support via a decision tree (multi-component) was better than either the workshop alone (ability) or decision tree alone (ability) in promoting appropriate prescriptions for patients with osteoarthritis.[112] The authors reported only significant comparisons of each of these intervention arms with a no-treatment control group and did not compare the intervention arms to each other. The authors suggested that the multi-component approach they used may confer a weak advantage compared with either of the ability strategies alone, but they presented no specific statistical analyses supporting this statement. The second cRCT with a 6-month follow-up period tested whether continuing medical education (CME) (ability strategy) that also provided practice enablers and reinforcers (motivation strategy) was better than CME alone in promoting guideline adherence.[106] Enablers and reinforcers took the form of a nurse providing prompts to reevaluate care for patients who might have undermanaged, high-risk cardiovascular disease and a checklist of the most recent and relevant clinical practice guidelines. The arm with the CME plus these reinforcing intervention elements was significantly better than CME alone. The RCT in this category was large (N=2,624). They compared an ability strategy (training sessions) with a multi-component strategy (training sessions plus provision of portable spirometry) to improve diagnosis and severity classification of chronic obstructive pulmonary disease.[114] The authors compared both arms with a no-treatment control group, but not to each other, and did not report confidence intervals or significance tests for our comparisons of interest.

Strategies to Increase Motivation Versus Multi-component Strategies

One trial tested a dissemination strategy to affect clinician behavior by increasing motivation against a multi-component strategy (Table 16);[96] it used a cRCT design with a 360-day follow-up. The primary outcome was guideline adherence related to lymph node assessment and biopsy. The authors compared a motivation strategy involving an expert opinion leader with a multi-component strategy that involved the expert opinion leader plus academic detailing and a toolkit. The multi-component arm was more effective for one of the outcomes, i.e., lymph node assessment; the groups did not differ significantly, however, for lymph node removal.

Strategies to Increase Reach and Ability Versus Multi-Component Strategies

One four-arm cRCT study compared a reach strategy (guidelines by mail), two different ability strategies (one educational session with guidelines and a continuing professional education session with guidelines), and a multi-component strategy that involved all these steps to affect the dispensing of antifungal medications for vulvovaginal candidiasis in community-based pharmacies (Table 16).[116] This trial assessed the primary outcome by having trained actors visit pharmacies presenting with particular clinical symptoms that might or might not require anti-fungal medication. The groups did not differ significantly immediately after the intervention.

Detailed Synthesis: Dissemination Strategies for Clinicians and Clinical Outcomes

Table 17 documents the strength of evidence grading for three trials that examined dissemination strategies focused on clinicians and assessed clinical outcomes and gives the overall SOE grade.

Table 17. Strength of evidence: Studies examining dissemination to clinicians for clinical outcomes

Strategy	Number of Trials; Subjects; Design/Quality	Risk of Bias	Consistency	Directness	Precision	Results and Strength of Evidence
Reach strategies versus ability strategies	1; N_c = 9[94] N patients=9,820 N clinicians=781 cRCT/Good	Low	Unknown (single study)	Direct	Precise	No significant differences among groups receiving reach or ability strategies, and confidence intervals or p values not reported for some comparisons Insufficient
Reach strategies versus multi-component strategies	4; N_c= 186 cRCT/Good[84,93,104] cRCT/Fair[101] cRCT/Good[104] cRCT/Good[93]	Low	Consistent	Direct	Imprecise	No significant differences in any trial between groups receiving reach or multi-component strategies. Low for no difference
Reach strategies, ability strategies, and Multi-component strategies	1; N_c = 18 cRCT/Fair[103]	Moderate	Unknown (single study)	Direct	Imprecise	No significant differences between groups receiving multi-component reach, or ability strategies. Insufficient

Abbreviations: cRCT = cluster randomized controlled trial; RCT = randomized controlled trial.

Table 18 describes clinician-focused trials and summarizes their results for clinical outcomes. Text accompanying the table describes trials that compare (a) reach and ability strategies, (b) reach and multi-component strategies, and (c) reach, ability, or multi-component strategies.

Table 18. Summary of trials examining dissemination to clinicians for clinical outcomes

Author, Year, Design Sample Size Duration Quality Grade	Setting Sample Characteristics	Intervention Groups	Outcome	Results
Trails that compared reach strategies with abilty strategies				
Simon et al., 2005[94] cRCT N_c = 9 N (clinicians) = 781 N (patients) = 9,820 Baseline, 1-year follow-up, 2-year follow-up Good	Clinical All patients with hypertension receiving primary care at 1 of 9 study sites; all clinicians providing primary care for adults at the 9 study sites	G1: Mailed educational materials (increase reach) G2: Individual academic detailing (increase ability) G3: Group academic detailing (increase ability)	Rates of hospitalization (2nd) Blood pressure control (2nd)	G1 = G2 G1=G3 No significant difference among groups for blood pressure control. Confidence intervals or p values not reported for G2 vs. G3 comparison for blood pressure control. Confidence intervals or p values not reported for any group comparisons for rates of hospitalizations.
Trails that compared reach strategies with multi-component strategies				
Bekkering et al., 2005[84] cRCT N_c= 68 N (physiotherapist) = 113 N (patient) =511 Baseline, 6, 12, 26, and 52 weeks after baseline Good	Clinical Physcial therapy practices that were members of the Royal Dutch Society for Physical Therapy (KNGF) located in or around the cities of Utrecht, Amersfoort, and Hilversum in the Netherlands	G1: Received guidelines by mail (increase reach) G2: Received guidelines + active training strategy (multi-component)	Patient physical functioning (P) Pain (P)	G2 = G3 No significant difference between groups.

Table 18. Summary of trials examining dissemination to clinicians for clinical outcomes (continued)

Author, Year, Design Sample Size Duration Quality Grade	Setting Sample Characteristics	Intervention Groups	Outcome	Results
Davis et al., 2004[101] cRCT N_c = 68 N= 2,025 Baseline, 12 month follow-up Fair	Clinical GP: All general practices in Tayside, UK Patient: Patients on the lists of participating GPs who were receiving medication for epilepsy and were older than 16 years	G1: Control - guidelines by mail (increase reach) G2: Intermediate (multi-component) G3: High intervention (multi-component)	SF-36 general health-related quality-of-life instrument (P)	G1 = G2 = G3 No significant difference between groups as reported by authors, no CI or p value reported for comparison
Jain et al., 2006[104] cRCT N_c = 50 Baseline, 12 month follow-up Good	Clinical Practice: ICU in Canada with at least 8 beds and a dietician Patient: All mechanically ventilated ICU patients who remained in the ICU for > 72 hours	G1: Passive intervention-guidelines by mail (incease reach) G2: Active intervention (multi-component)	Nutritional adequacy of enteral nutrition (P) Glycemic control (2nd) Duration of ICU or hospital stay (2nd) 28 day mortality (2nd)	G1 = G2 No significant difference between groups for nutritional adequacy, duration of ICU or hospital stay, and 28 day mortality, although p values or CI were only reported for nutritional adequacy. G2>G1 for 3 measures of glycemic control
Rebbeck et al. 2006[93] cRCT N_c = 27 N = 99 Baseline, 1.5 months, 3 months, 6 months, and 12 months Good	Clinical Physiotherapy clinics in Australia Clinic: Clinics in 2 states in Australia that had seen at least 5 whiplash cases in the previous year Patient: Patients, 18 years and older, involved in a motor vehicle accident within the previous 6 weeks who presented to the clinic with acute whiplash	G1: Dissemination of guidelines by mail (increase reach) G2: Implementation group (multi-component)	Patient disability (P)	G1 = G2 No significant difference between groups.

Table 18. Summary of trials examining dissemination to clinicians for clinical outcomes (continued)

Author, Year, Design Sample Size Duration Quality Grade	Setting Sample Characteristics	Intervention Groups	Outcome	Results
Trails that compared reach strategies, ability strategies and multi-component strategies				
Hagmolen et al., 2008[103] cRCT N_c = 18 Baseline, 12-month follow-up Fair	Clinical Children ages 7 to 17 years, who had at least two prescriptions of β2-agonists or ICS in the year before invitation	G1: Guideline dissemination (increase reach) G2: Guideline dissemination + educational program (increase ability) G3: Guideline dissemination + educational program + individualized treatment advice based on airway responsiveness and symptoms (multi-component)	Change in airway hyper-resonsiveness in children after 1 year (P) Change in asthma symptom scores (2nd) • Total sypmtoms score • Noctural symptoms score Usage of asthma medication (2nd): • Inhaled corticosteroids • B_2-agonist	G1 = G2 = G3 No significant differences among groups total asthma symptom score, and use of B_2-agonist, G1 and G3>G2 for nocturnal asthma symptoms with largest improvement in G1 G3>G1 and G2 for Inhaled corticosteroids

Note: In most cases, the labels of the intervention groups represent the labels that study authors had originally used, except if they were merely labeled control or intervention. Information in parentheses represents the strategy that the authors of this review believed best captured the intervention components. When authors used nondescriptive or vague labels for the intervention groups (e.g., Group 1, Intervention A) we used a summary of their description of the group as a label. For a full description of the studies intervention components refer to Appendix E.

Abbreviations: 2nd = secondary outcome; + = Plus; COPD = chronic obstructive pulmonary disease; cRCT = cluster randomized controlled trial; GP = general practitioners; ICS = inhaled corticosteroids; ICU = intensive care unit; P = primary outcome; RCT = randomized controlled trial; SEPAR = Sociedad Española de Neumología y Cirugía Torácica.

Strategies to Increase Reach Versus Strategies to Increase Ability

One cRCT compared a reach strategy with two ability strategies.[94] The reach strategy involved mailed information. The ability strategies involved group academic detailing in one group and individual academic detailing in the other group. Over a two year follow-up period there were no significant differences among groups for guideline adherence related to treatment of hypertension patients.

Strategies to Increase Reach Versus Multi-component Strategies

Four studies compared reach strategies with multi-component strategies.[84,93,101,104] Three were cRCTs that had 1-year follow-up assessing patient outcomes such as physical functioning and pain,[84] nutrition support,[104] or quality of life.[101] One study was an cRCT with follow-up at 1.5, 3, 6, and 12 months.[93] Three studies used mailed guidelines as the reach strategy; multi-component approaches involved educational training, use of supportive decision aids, opinion leaders, or a nurse specialist for consultation.[84,93,101] Jain et al. compared giving written materials with an extensive multi-component intervention providing education, web-based tools, and interpersonal support to dieticians providing nutritional support to critically ill adults on mechanical ventilation in an intensive care unit.[104] The reach and multi-component approaches did not differ significantly for any of these studies at follow-up assessment.

Strategies to Increase Reach, Ability, and Multi-Component Strategies

One trial tested the effect of dissemination strategies to improve clinical outcomes by increasing reach versus increasing ability versus a multi-component approach.[103] Reach strategy included dissemination of a guideline regarding childhood asthma treatment, ability strategy included the addition of education, and the multi-component approach added individualized treatment advice. Study groups did not differ significantly in in airway responsiveness at 12-month follow-up.

Detailed Synthesis: Dissemination Strategies for Clinicians and Knowledge Outcomes

Table 19 documents the strength of evidence grading for each of the dissemination comparisons focused on clinicians for clinical outcomes and gives the overall SOE grade.

Table 19. Strength of evidence: Trials examining dissemination to clinicians for knowledge outcomes.

Strategy	Number of Studies; Subjects; Design/Quality Rating	Risk of Bias	Consistency	Directness	Precision	Results and Strength of Evidence
Reach strategies versus ability strategies	1; N = 213 RCT/Fair[115]	Moderate	Unknown (single study)	Direct	Imprecise	Ability strategy significantly more effective than reach strategy Insufficient
Reach strategies versus multi-component strategies	1; N_c = 27 N = 99 cRCT/Good[93]	Low	Unknown (single study)	Direct	Imprecise	Multi-component strategy significantly more effective than reach strategy Insufficient
Reach strategies, ability strategies, and multi-component strategies	1; Nc = 61 cRCT/Fair[116]	Moderate	Unknown (single study)	Direct	Imprecise	No significant differences between groups receiving reach, ability, and multi-component approaches Insufficient

Abbreviations: cRCT = cluster randomized controlled trial; RCT = randomized controlled trial.

Table 20 describes clinician-focused trials and summarizes their results for knowledge outcomes. Text accompanying the table describes trials that compare (a) reach and ability strategies, (b) reach and multi-component strategies, and (c) reach, ability, or multi-component strategies.

Strategies to Increase Reach versus Strategies to Increase Ability

One trial compared a reach strategy (dissemination of guidelines via email) with an ability strategy (guidelines and a web-based education program) and measured enhancement of residents' knowledge of the use of opioids for managing pain.[115]At the end of a 60-day follow-up period, residents receiving the ability strategy had significantly higher scores on a knowledge and competence measure.

Strategies to Increase Reach versus Multi-Component Strategies

One study compared a reach strategy (dissemination of guidelines via mail) with a multi-component strategy that involved a one-day workshop led by opinion leaders. During the workshop materials supporting guideline practice for whiplash were provided to participants. The workshop was followed by a two-hour outreach visit six months later. The multi-component approach was significantly better at increasing knowledge about how to treat whiplash.

Table 20. Summary of studies examining dissemination to clinicians for knowledge outcomes

Author, Year, Design Sample Size Duration Quality Rating	Setting Sample Characteristics	Intervention Groups	Outcome	Results
Trials that compared reach strategies with ability strategies				
Sullivan et al., 2010[115] RCT N = 213 Baseline, immediate posttest, 60-day posttest Fair	Clinical Residents in internal medicine	G1: VA guidelines for pain management (increase reach) G2: COPE: web-based education program for pain management (increase ability)	Residents' knowledge and competence with managing opioids for chronic non-cancer pain (P)	G2 > G1 Group with interactive web-based training significantly better knowledge scores than the group with standard guidelines.
Trials that compared reach strategies with multicomponent strategies				
Rebbeck et al. 2006[93] cRCT N_c = 27 N = 99 Baseline, 1.5 months, 3 months, 6 months, and 12 months Good	Clinical Physiotherapy clinics in Australia Clinic: Clinics in 2 states in Australia that had seen at least 5 whiplash cases in the previous year Patient: Patients, 18 years and older, involved in a motor vehicle accident within the previous 6 weeks who presented to the clinic with acute whiplash	G1: Dissemination of guidelines by mail (increase reach) G2: Implementation group (multi-component)	Knowledge about evidence (P)	G2 > G1 Group with multicomponent strategy significantly better knowledge scores than group with standard dissemination.

Table 20. Summary of studies examining dissemination to clinicians for knowledge outcomes (continued)

Author, Year, Design Sample Size Duration Quality Rating	Setting Sample Characteristics	Intervention Groups	Outcome	Results
Trials that compared reach strategies, ability strategies, and multi-component strategies				
Watson et al., 2002[116] cRCT N_c = 61 Baseline, immediate posttest Fair	Community-based settings Community pharmacies in Scotland	G1: Guideline materials by postal mail (increase reach) G2: One education outreach (EO) session and guidelines (increase ability) G3: Continuing professional education (CPE) session (increase ability) G4: Gudielines + EO + CPE (multi-component)	Pharmacists' knowledge of the treatment of vaginal candidiasis (P)	G1 = G2 = G3 = G4 No significant differences among groups as reported by authors. Confidence intervals or p values not reported for comparison.

Note: In most cases, the labels of the intervention groups represent the labels that study authors had originally used, except if they were merely labeled control or intervention. Information in parentheses represents the strategy that the authors of this review believed best captured the intervention components. When authors used nondescriptive or vague labels for the intervention groups (e.g., Group 1, Intervention A) we used a summary of their description of the group as a label. For a full description of the studies intervention components refer to Appendix E.

Abbreviations: COPE = Collaborative Opioid Prescribing Education; cRCT = cluster randomized controlled trial; P = primary outcome; RCT = randomized controlled trial; VA = Department of Veterans Affairs.

Strategies to Increase Reach, Strategies to Increase Ability, and Multi-Component Strategies

This four-arm cRCT study (previously described in Table 16) also assessed clinician knowledge as a primary outcome.[116] They compared a reach strategy (guidelines by mail), two ability strategies (educational session with guidelines and CME with guidelines), and a multi-component strategy that involved all these combined strategies. The outcome measure appropriate dispensing of antifungal medications for vulvovaginal candidiasis in community-based pharmacies. The groups did not differ significantly in knowledge scores assessed with self-report pre-post intervention questionnaires

Key Question 2: Disseminating Evidence to Patients

Key Points:

- Evidence is inconsistent for determining the benefit of reach, ability, motivation, or multi-component approaches for patients focused on changing health-related decisions and behaviors. (12 trials; insufficient SOE).
- Evidence is insufficient for determining the benefit of reach, ability, motivation or multi-component approaches for patients focused on changing clinical outcomes. (2 trials; insufficient SOE due to one trial in each category).
- Evidence is insufficient for determining the benefit of reach, ability, motivation or multi-component approaches for patients focused on changing knowledge outcomes (3 trials; insufficient SOE due to inconsistent findings or one trial in a category).

Detailed Synthesis: Dissemination Strategies for Patients and Health-Related Decisions or Behaviors

Table 21 documents the strength of evidence grading for each of the dissemination comparisons focused on patients for health-related decisions and behavior outcomes and gives the overall SOE grade.

Table 21. Strength of evidence: Trials examining dissemination to patients for health-related decisions and behavior outcomes.

Strategy	Number of Studies; Subjects; Design/Quality Rating	Risk of Bias	Consistency	Directness	Precision	Results and Strength of Evidence
Reach strategies of various sorts	3; N= 1,710 RCT/Fair[102,110,111]	Moderate	Inconsistent	Direct	Precise	Inconsistent effects of comparisons for various reach strategies Insufficient
Reach strategies versus motivation strategies	4; N=1,609 3 RCT/Fair[99,107,118] 1 RCT/Good[89]	Moderate	Inconsistent	Direct	Precise	Inconsistent effects of reach and motivation strategies on mammography screening adherence and physical activity. Insufficient
Reach strategies versus multi-components strategies	4; N= 2,591 2 RCT/Good[90,119] 2 RCT/Fair[80,113]	Moderate	Inconsistent	Direct	Precise	Inconsistent effects of reach and multi-component strategies on dietary behaviors, mammography or smoking cessation. Insufficient
Motivation strategies versus multi-component strategies	1; N=287 fRCT/Fair[98]	Moderate	Unknown (single study)	Direct	Imprecise	No results reported for comparisons between active comparators. Insufficient

Abbreviations: fRCT = factorial randomized controlled trial; RCT = randomized controlled trial.

Table 22 describes patient-focused trials and summarizes their results for health-related decisions and behavior outcomes. Text accompanying the table describes trials that (a) compare reach strategies of various types, (b) reach and motivation strategies, (c) reach and multi-component strategies, and (d) motivation and multi-component strategies.

Strategies to Increase Reach

Three trials compared reach strategies with other reach strategies with respect to patient health-related decisions and behaviors including intention to get a PSA screening,[102,111] self-efficacy for infant care, health care utilization,[110] and evidence discussions.[111] These RCTs had samples sizes ranging from 137 to 1,152, took place in community-based, academic heatlth care, and outpatient clinical settings. Follow-up assessment ranged from 1 week to 2 months after the interventions. Results across the trials were inconsistent, but generally did not find any significant differenences between the reach strategies compared. The two studies that compared printed materials withelectronic methods—i.e., a DVD[110] or a video[111] found no significant differences between groups on outcomes, except that the DVD was significantly more effective at reducing two measures of health care utilization[110] among parents caring for infants.

Table 22. Summary of trials examining dissemination to patients for health-related decisions and behavior outcomes

Author, Year, Design Sample Size Duration Quality Rating	Setting Sample Characteristics	Intervention Groups	Outcome	Results
Trials that compare reach strategies				
Gattellari and Ward, 2005[102] RCT N = 421 21 days median length between pre-test and post-test Fair	Community-based settings Men ages 50 to 70 with no known history of prostate cancer	G1: Leaflet (increase reach) G2: Video (increase reach) G3: Booklet (increase reach)	Intention to get PSA screening in next year (2nd)	G1 = G2 = G3 No significant differences between groups.
Paradis et al., 2011[110] RCT N = 137 Baseline, 2 weeks and 2 months post-intervention Fair	Academic health care institutions Parents or primary caregivers over18 years old of newborns less than 1 month old presenting for their first visit to the pediatrician's office.	G1: Paper handouts (increase reach) G2: Educational DVD (increase reach)	Self-efficacy for infant care skills (P) Five health care utilization measures (2nd)	G1 = G2 No significant difference between groups for self-efficacy. Mixed findings for utilization measures: G2>G1 Educational DVD more effective at reducing professional consultations and office visits G1=G2 No significant differences for additional clinic visits, parent phone calls to clinic, and emergency department visits
Partin et al, 2004[111] RCT N = 1152 1 week post Fair	Clinical Male veterans ages 50 and older who had no prostate cancer and a scheduled primary care appointment at 1 of 4 Veteran Affairs facilities in the Midwest	G1: Usual care (not abstracted) G2: Pamphlet (increase reach) G3: Video (increase reach)	Discussion about the evidence (2nd) Behavioral intention to get PSA test in next year (2nd) Getting PSA screening (2nd)	Confidence intervals or p values not reported for comparison for these outcomes

Table 22. Summary of trials examining dissemination to patients for health-related decisions and behavior outcomes (continued)

Author, Year, Design Sample Size Duration Quality Rating	Setting Sample Characteristics	Intervention Groups	Outcome	Results
Trials that compare reach strategies with motivation strategies				
Carney et al., 2005[99] RCT N = 258 Baseline, 12 and 27 month follow-up Fair	Community-based settings Women ages 50 and older, without a history of breast cancer, who were eligible for routine mammogram screening	G1: Mailed health information (increase reach) G2: Telephone counseling (increase motivation)	Adherence to mammography screening (P) Time between screening exams (2nd)	G2 > G1 Tailored telephone counseling significantly more effective in the short term (up to1 year post-intervention) than mailed guideline for adherence to mamography screening and time between screening G2>G1 Tailored telephone counseling significantly more effective in the long term (up to 27 months post-intervention) than mailed guideline for time between screenings
Kennedy et al., 2003[89] RCT N = 894 Baseline (6-weeks preconsultation), immediate post-consulation, 6, 12, and 24 months post-consultation Good	Clinical Women referrred to 1 of 28 consultant gynecologists from 6hospitals in England with a new episode of menorrhagia	G1: Control (not abstracted) G2: Information (increase reach) G3: Interview (increase motivation)	Behavioral intentions (treatment preferences) (2nd) Treatments undertaken for menorrhagia (2nd)	Confidence intervals or p values not reported for comparison for behavioral intentions G2 = G3 No significant difference between groups for treatment undertaken

Table 22. Summary of trials examining dissemination to patients for health-related decisions and behavior outcomes (continued)

Author, Year, Design Sample Size Duration Quality Rating	Setting Sample Characteristics	Intervention Groups	Outcome	Results
King et al., 2007[118] RCT N = 218 Baseline, 6, and 12 months Fair	Community-based settings Inactive men and women ages 55 and older	G1: Attention control (not abstracted) G2: Counselor via phone (increase motivation) G3: Automated counselor via phone (increase reach)	Physical activity measures: (P) • Stanford 7-day physical activity recall • CHAMPS	G2 = G3 No significant difference between groups.
Marcus et al., 2009[107] RCT N = 239 Baseline, 6 months and 12 months Fair	Community-based settings Sedentary, healthy adults ages 18 to 65	G1: Contact control treatment delayed group (not abstracted) G2: Telephone-based individualized feedback (increase motivation) G3: Print-based individualized feedback (increase reach)	7-day physical activity recall: (P) • Behavioral intentions to exercise (decisional balance) (2nd) • Self-efficacy for exercising (2nd)	G3 > G2 Print-based feedback significantly more effective than telephone-based feedback for physical activity and self-efficacy at 12 months but not at 6 months G3=G2 Print-based feedback not significantly more effective than telephone-based feedback for behavioral intentions at any time point

Table 22. Summary of trials examining dissemination to patients for health-related decisions and behavior outcomes (continued)

Author, Year, Design Sample Size Duration Quality Rating	Setting Sample Characteristics	Intervention Groups	Outcome	Results
Trials that compared reach strategies with multi-components strategies				
Elder et al., 2005[80] RCT N = 357 Baseline, 12-week follow-up Fair	Community-based settings Spanish-language-dominant women between 18 and 65 years	G1: Culturally targeted print-materials + activity inserts (increase reach) G2: Tailored print materials + activity inserts + supporting materials (multi-component). G3: Tailored print materials + in-person promotora (multi-component)	Percent calories from fat (P) Number of daily grams of fiber (P) Nine dietary variables (2nd): • Energy intake • Total fat • Total saturated fat • Soluble dietary fiber • Insoluble dietary fiber • Total carbohydrates • Glucose • Fructose • Sucrose	G1 = G2 = G3 No significant differences among groups G3>G1 for energy intake and total carbohydrates Tailored print materials plus promotora more effective than culturally targed materials in decreasing energy intake and total carbohydrates G3>G2 for total fat, total saturated fat, glucose and fructose Tailored print materials plus promotora more effective than tailored print materials in decreasing total fat, total saturated fat, glucose and fructose intake.
Lien et al., 2007[90a] Svetkey et al., 2003[92]Young et al., 2009[91] RCT (single trial; different outcomes) Baseline, 3, 6, 12 and 18 months N = 810 Good	Clinical Generally health adults over 25 years old with prehypertension or stage-1 hypertension	G1: Advice only (increase reach) G2: Advice + behavioral counseling using established intervention (multi-component) G3: Established intervention + DASH dietary recommendations (multi-component)	Meeting at least 3 dietary and fitness goals (P)	G3>G2>G1 Both strategies involving behavioral counseling significantly more effective than the individual advice without behavioral counseling at 6 months. G1=G2; G3>G2; G3 >G1 Multi-component approach involving DASH diet and behavioral counseling significantly more effective than the individual advice without behavioral counseling at 18 month and behavioral counseling with established intervention.

Table 22. Summary of trials examining dissemination to patients for health-related decisions and behavior outcomes (continued)

Author, Year, Design Sample Size Duration Quality Rating	Setting Sample Characteristics	Intervention Groups	Outcome	Results
Rimer et al., 2001[113] RCT N = 1,127 Yearly Fair	Clinical Women in their mid-40s and mid-50s	G1: No treatment control/usual care (not abstracted) G2: Tailored print (TP) (increase reach) G3: TP + telephone counselling (TC) (multi-component)	Adherence to yearly screening mammography (P)	G2 = G3 No significant difference between groups.
Wetter et al., 2006[119] RCT N = 297 Baseline, 5-week follow-up, 12-week follow-up Good	Community-based settings Spanish-speaking smokers, ages 18 and older, who called the National Cancer Institute Cancer Information Service to request smoking cessation	G1: Single standard telephone-counseling session (increase reach) G2: Multiple enhanced telephone counseling sessions (multi-component)	Smoking abstinence (P)	G2 > G1 Proactive (enhanced, multiple sessions) telephone counseling plus culturally tailored information significantly more effective than standard single session) telephone-counseling.
Trials that compared motivation strategies with multi-component strategies				
Campbell et al., 2004[98] fRCT N = 287 Baseline , 12-month follow-up Fair	Community-based settings Adults, age 18 and older, that were active members of participating church	G1: Control (not abstracted) G2: Lay health advsior (LHA) (increase motivation) G3: Tailored and targeted print and video (TPV) (multi-component) G4: TPV and LHA (multi-component)	Fruit and vegetable consumption (P) Physical activity (P) Fat intake (P) CRC screening (2nd)	G3>G2, G3>G4 Tailored print and video significantly more effective than LHA or multi-component approach for increasing fruit and vegetable intake and physical activity G2=G3=G4 No significant differences for CRC screening and fat intake

Note: In most cases, the labels of the intervention groups represent the labels that study authors had originally used, except if they were merely labeled control or intervention. Information in parentheses represents the strategy that the authors of this review believed best captured the intervention components. When authors used nondescriptive or vague labels for the intervention groups (e.g., Group 1, Intervention A) we used a summary of their description of the group as a label. For a full description of the studies intervention components refer to Appendix E.

Abbreviations: 2nd = secondary outcome; + = plus; CRC = colorectal cancer; P = primary outcome.

Strategies to Increase Reach Versus Strategies to Increase Motivation

Four trials (N ranging from 218 to 894) compared strategies to enhance reach with strategies to enhance motivation. All trials examined findings over a 6 and 12-month follow-up periods, and the two used 24 or 27-month follow-up periods. One aimed to encourage adherence to mammography screening guidelines;[99] two encouraged physical activity among older adults in community settings,[107,118] and the other examined treatment preferences and treatments undertaken for menorrhagia.[89] Interpersonal telephone counseling, categorized as a motivational strategy, produced inconsisent effects. In one trial, telephone counseling was significantly better than a reach strategy involving mailed information, at prompting recipients to get mammography screening.[99] In another trial, interpersonal telephone counseling and automated telephone counseling did not differ in encouraging physical activity over 1 year[118]; in a third trial, print-based feedback was more effective than telephone-based counseling in promoting physical activity over 1 year at at the 12-, but not 6-month, follow-up.[107] In the trial that compared provision of an information packet (reach) to a preference elicitation interview with a nurse (motivation) found no differences for the outcomes studied.[89]

Strategies to Increase Reach Versus Multi-Component Strategies

Four trials (N ranging from 297 to 1,127) compared reach strategies with multi-component approaches. Two took place in community settings and focused on enhancing Latino health, with one focused on changing nutrition intake[80] and one focused on smoking cessation.[119] Both trials had a 12-week follow-up. Another trial also focused on increasing dietary and fitness goals.[90] The other trial, in a clinical setting, focused on improving adherence to mammography screening.[113] Two of these trials compared reach strategies (print materials) with a multi-component strategy (print materials plus interpersonal counseling via telephone or in person); groups did not differ significantly in the primary outcomes studied, percent carlories from fat or number of daily grams of fiber,[80,113] and of the nine secondary outcomes related to dietary intake only three were significantly increased in the group that received tailored print materials plus an interpersonal counselor.[80] The trial that compared single-session telephone counseling (motivation strategy) with more intensive telephone counseling plus supporting materials (multi-component) and found the multi-component strategy to be more effective than the motivation strategy at the 12-week follow-up.[119] The trail that compared the provision of advice and information to two differenent behavioral counseling approaches that differed by the type of dietary recommendations provided found inconsistent results.[90] Both multi-component strategies involving behahavioral counseling were more effective than just the provision of individual advise at a six month follow-up. However, at an 18-month follow-up the multi-component strategy involving the DASH dietary recommendations was signficantly more effective than the other groups.

Strategies to Increase Motivation Versus Multi-Component Strategies

One trial compared an interpersonal motivation strategy (lay health advisors [LHA]) with two different multi-component approaches--one that involved tailored and targeted print and video materials (TPV), and another that involved LHAs plus the print and video materials (LHA+TPV).[98] TPV was significantly better than control, and a LHA or a multi-component approach that involved both TPV and LHA for fruit and vegetable intake and physical activity; the more intense multi-component approach (LHA+TPV) was not significantly better than

control or other active comparators for any outcome. There were no significant differences among groups related to fat intake or CRC screening.

Detailed Synthesis: Dissemination Strategies to Patients for Clinical Outcomes

Table 23 documents the strength of evidence grading for each of the dissemination comparisons focused on patients and assessing clinical outcomes and gives the overall SOE grade.

Table 23. Strength of evidence: Trials examining dissemination to patients for clinical outcomes

Strategy	Number of Studies; Subjects; Design/Quality Rating	Risk of Bias	Consistency	Directness	Precision	Results and Strength of Evidence
Reach strategies versus motivation strategies	1; N= 894 RCT/Good[89]	Low	Unknown (single study)	Direct	Precise	No significant differences between groups. Insufficient
Reach strategies versus multi-component strategies	1[a]; N= 810 RCT/Good[90-92]	Low	Inconsistent (single trial population but different outcomes)	Direct	Precise	Inconsistent results for reach and multi-component groups. Insufficient

[a] These articles reported on the same sample population but reported slightly different outcomes. We combined information from these articles and report them as one article in text.

Abbreviations: N = number; RCT = randomized controlled trials

Table 24 describes patient-focused trials and summarizes their results for clinical outcomes. Text accompanying the table describes trials that compare (a) reach and motivation strategies, and (b) reach and multi-component strategies.

Strategies to Increase Reach versus Strategies to Increase Motivation

One trial compared a reach strategy that involved providing written and video-based information with a preference elicitation interview for 894 women with menorrhagia in outpatient clinics.[89] Follow-up measures at 6, 12, and 24 months showed no differences between groups on a health status measure.

Strategies to IncreaseReach versus Multi-Component Strategies

One trial compared a reach strategy involving providing advice in an in-person meeting with registered dietitian with two different multi-component strategies that entailed lifestyle behavioral counseling with slightly different informational content.[90-92] Participants were adults with or at risk for hypertension. The two multi-component groups did not differ significantly at 6-month follow-up, but both were superior to advice only in affecting blood pressure at 6-month follow-up.

Table 24. Summary of trails examining dissemination to patients for clinical outcomes

Author, Year, Design Sample Size Duration Quality Grade	Setting Sample Characteristics	Intervention Groups	Outcome	Results
Trials that compared reach strategies with motivation strategies				
Kennedy et al., 2003[89] RCT N = 894 Baseline (6-weeks preconsultation), immediate post-consulation, 6, 12, and 24 months post-consultation Good	Clinical Women referrred to 1 of 28 consultant gynecologists from 6hospitals in England with a new episode of menorrhagia	G1: Control (not abstracted) G2: Information (increase reach) G3: Interview (increase motivation)	Health-related quality of life measured by the Medical Outcomes Study Short Form (SF-36) questionnaire (P)	G2 = G3 No significant difference between groups.

Table 24. Summary of trails examining dissemination to patients for clinical outcomes (continued)

Author, Year, Design Sample Size Duration Quality Grade	Setting Sample Characteristics	Intervention Groups	Outcome	Results
Trials that compared reach strategies with multi-component strategies				
Lien et al., 2007[90a] Svetkey et al., 2003[92]Young et al., 2009[91] RCT (single trial; different outcomes) Baseline, 3, 6, 12 and 18 months N = 810 Good	Clinical Generally health adults over 25 years old with prehypertension or stage-1 hypertension	G1: Advice only (increase reach) G2: Advice + behavioral counseling using established intervention (multi-component) G3: Established intervention + DASH dietary recommendations (multi-component)	Change in systolic blood pressure at 6 months (P) Change in weight (2nd)	G2 > G1; G3 > G1 G3 = G2 Both strategies involving behavioral counseling significantly more effective than the individual advice without behavioral counseling for reducing blood pressure and weight.

Note: In most cases, the labels of the intervention groups represent the labels that study authors had originally used, except if they were merely labeled control or intervention. Information in parentheses represents the strategy that the authors of this review believed best captured the intervention components. When authors used nondescriptive or vague labels for the intervention groups (e.g., Group 1, Intervention A) we used a summary of their description of the group as a label. For a full description of the studies intervention components refer to Appendix E.

Abbreviations: 2nd = secondary outcome; DASH = Dietary Approaches to Stop Hypertension; P = primary outcome.

Detailed Synthesis: Dissemination Strategies for Patients and Knowledge Outcomes

Table 25 documents the strength of evidence grading for each of the dissemination comparisons focused on patients for knowledge outcomes and gives the overall SOE grade.

Table 25. Strength of evidence: Trails examining dissemination to patients for knowledge outcomes

Strategy	Number of Studies; Subjects; Design/Quality; Rating	Risk of Bias	Consistency	Directness	Precision	Results and Strength of Evidence
Reach strategies of various sorts	3; N= 1,710 RCT/Fair[102,110,111]	Moderate	Inconsistent	Indirect	Precise	Inconsistent findings about the benefit of print or video reach strategies. Insufficient
Reach strategies versus motivation strategies	1; N =894 RCT/Good	Low	Unknown (single study)	Indirect	Precise	Confidence intervals or p values not reported for this comparison Insufficient
Reach strategies versus multi-component strategies	1; N= 1,127 RCT/Fair[113]	Moderate	Unknown (single study)	Indirect	Precise	Inconsistent findings across two knowledge outcomes. Insufficient

Table 26 describes patient-focused trials and summarizes their results for knowledge outcomes. Text accompanying the table describes trials that compare (a) various reach strategies. (b) reach and motivation strategies, and (c) reach and multi-component strategies.

Strategies to Increase Reach

Three trials compared reach strategies with other reach strategies with respect to patient knowledge. These RCTs had samples sizes ranging from 137 to 1,152, took place in community-based, academic heatlth care, and outpatient clinical settings. Follow-up assessment ranged from 1 week to 2 months after the interventions. Two studies related to prostate cancer screening knowledge; the other focused on parental knowledge of infant development. Results across the trials were inconsistent. The two studies that compared printed materials withelectronic methods—i.e., a DVD[110] or a video[111] found no significant differences between groups on knowledge outcomes. The two studies that examined prostate cancer screening knowledge showed inconsistent results.[102,111] In one, a decision aid in the form a booklet was significantly more effective in increasing knowledge than either a video or leaflets;[102] by contrast, groups receiving written materials or a video did not differ significantly (although these were not designed as decision aids).[111]

Table 26. Summary of trails examining dissemination to patients for knowledge outcomes

Author, Year, Design Sample Size Duration Quality Rating	Setting Sample Characteristics	Intervention Groups	Outcome	Results
Trials that compared reach strategies				
Gattellari and Ward, 2005[102] RCT N = 421 21 days median length between pre-test and post-test Fair	Community-based settings Men ages 50 to 70 with no known history of prostate cancer	G1: Leaflet (increase reach) G2: Video (increase reach) G3: Booklet (increase reach)	Knowledge about prostate cancer screening (P)	G3 > G2, G3 > G1 G2 = G1 Booklet significantl more effective than either leaflet or video.
Paradis et al., 2011[110] RCT N = 137 Baseline, 2 weeks and 2 months post-intervention Fair	Academic health care institutions Parents or primary caregivers over18 years old of newborns less than 1 month old presenting for their first visit to the pediatrician's office.	G1: Paper handouts (increase reach) G2: Educational DVD (increase reach)	Parent knowledge of infant development (P)	G1 = G2 No significant difference between groups.
Partin et al, 2004[111] RCT N = 1152 1 week post Fair	Clinical Male veterans ages 50 and older who had no prostate cancer and a scheduled primary care appointment at 1 of 4 Veteran Affairs facilities in the Midwest	G1: Usual care (not abstracted) G2: Pamphlet (increase reach) G3: Video (increase reach)	Prostate Cancer screening knowledge (P)	G2 = G3 No significant difference between groups as reported by authors, no significance tests or CI reported.

Table 26. Summary of trails examining dissemination to patients for knowledge outcomes (continued)

Author, Year, Design Sample Size Duration Quality Rating	Setting Sample Characteristics	Intervention Groups	Outcome	Results
Trials that compared reach strategies with motivation strategies				
Kennedy et al., 2003[89] RCT N = 894 Baseline (6-weeks preconsultation), immediate post-consulation, 6, 12, and 24 months post-consultation Good	Clinical Women referrred to 1 of 28 consultant gynecologists from 6hospitals in England with a new episode of menorrhagia	G1: Control (not abstracted) G2: Information (increase reach) G3: Interview (increase motivation)	Knowledge of available treatment options (2nd)	Confidence intervals or p values not reported for this outcome

Table 26. Summary of trails examining dissemination to patients for knowledge outcomes (continued)

Author, Year, Design Sample Size Duration Quality Rating	Setting Sample Characteristics	Intervention Groups	Outcome	Results
Trials that compared reach strategies with multi-component strategies				
Rimer et al., 2001[113] RCT N = 1127 Yearly Fair	Clinical Women in their mid 40s and mid 50s	G1: No treatment control/usual care (not abstracted) G2: Tailored print (increase reach) G3: Tailored print + telephone counselling (multi-component)	Knowledge for: • Accuracy of risk perception (P) • Mammogram effectiveness (P)	Inconsistent findings G3 > G2 for risk perception accuracy; G3 = G2 for mammogram effectiveness Multi-component strategy was more effective than tailored print intervention for increasing knowledge about risk for cancer than the reach strategy, but the stragegies did not differ in knowledge about mammogram effectiveness.

Note: In most cases, the labels of the intervention groups represent the labels that study authors had originally used, except if they were merely labeled control or intervention. Information in parentheses represents the strategy that the authors of this review believed best captured the intervention components. When authors used nondescriptive or vague labels for the intervention groups (e.g., Group 1, Intervention A) we used a summary of their description of the group as a label. For a full description of the studies intervention components refer to Appendix E.

Abbreviations: 2nd = secondary outcome; P = primary outcome; RCT = randomized controlled trial.

Strategies to Increase Reach Versus Strategies to Increase Motivation

One trial (N = 894) compared a strategy to enhance reach with a strategy to enhance motivation using a 6, 12, and 24-month follow-up period. This trial did not report significance tests or confidence intervals for the group comparisons examining knowledge of treatment options for menorrhagia.[89]

Strategies to Increase Reach versus Multi-component Strategies

One trial (described in Table 22 for behavioral outcomes) also compared a reach strategy that used print materials with a multi-component strategy that used print materials plus interpersonal counseling by telephone to increase two types of knowledge—accuracy of risk perceptions and mammography effectiveness.[113] For knowledge outcomes, results were inconsistent. The multi-component approach was more effective than the reach strategy for accuracy of risk perceptions but not for mammography effectiveness.

Key Question 2: Disseminating Information to Clinicians and Patients

Key Points

- Evidence is inconsistent for determining the benefit of reach, ability, motivation, or multi-component strategies that target both providers and patients for health-related decisions and behaviors (4 trials; insufficient SOE).
- Evidence is inconsistent for determining the benefit of reach, ability, motivation, or multi-component strategies that target both providers and patients for health-related decisions and behaviors or clinical outcomes (1 trial in each category; insufficient SOE).

Detailed Synthesis: Dissemination Strategies for Clinicians and Patients for Health-related Decisions and Behavior Outcomes

Table 27 documents the strength of evidence grading for each of the dissemination comparisons focused on clinicians and patients for health-related decisions and behavior outcomes and gives the overall SOE grade.

Table 27. Strength of evidence: Trials examining dissemination to clinicians and patients for health-related decisions and behavior outcomes

Strategy	Number of Studies; Subjects; Design/Quality	Risk of Bias	Consistency	Directness	Precision	Results and Strength of Evidence
Reach strategies versus multi-component strategies	4; N= 1,676 RCT/Good[85] fRCT/Fair[100] RCT/Good[87] 1cRCT/Good[82] Nc=116,N=1378	Low	Inconsistent	Direct	Precise	Results inconsistent or authors did not present significance tests or confidence intervals for comparisons Insufficient
Ability strategies versus multi-component strategies	1; Nc=30 cRCT/Good[86]	Low	Unknown (single study)	Direct	Imprecise	No significant differences between groups Insufficient
Studies that compared Multi-component strategies	1; N_c = 548 cRCT/Fair[108]	Moderate	Unknown (single study)	Direct	Imprecise	Authors did not report significant tests for comparisons Insufficient

Table 28 describes clinician and patient-focused trials and summarizes their results for health-related decisions and behavior outcomes. Text accompanying the table describes trials that compare (a) reach and multi-component strategies, (b) ability and multi-component strategies, and (c) various multi-component strategies.

Strategies to Increase Reach Versus Multi-component Strategies

Three RCTs and one cRCT compared reach strategies with multi-component strategies to affect health-related decisions and behavior outcomes. Sample sizes ranged from 327 to 1387; and follow-up periods ranged from 12 weeks to 1 year. Reach strategies included dissemination to either a patient[100] or clinicians.[82,85,87] The multi-components strategies involved dissemination to both patients and physicians. Across these four trails there was either no significant difference between groups or significance test or confidence intervals were not reported for comparisons between active comparators.

Strategies to Increase Ability Versus Multi-component strategies

One cRCT study compared an academic detailing session (increase clinician ability) with a multi-component strategy that included academic detailing and tools and resources for both patients and providers.[86] Patient resources included a patient education toolkit (with a companion Web site) and a computer kiosk with patient activation software. Physicians received with a decision support tool based on a personal digital assistant, which included four booster academic detailing sessions. The clinician only and multi-component group focused on both clinicians and patients did not differ significantly.

Table 28. Summary of studies examining dissemination to clinicians and patients for health-related decisions and behavior outcomes

Author, Year, Design Sample Size Duration Quality Rating	Setting Sample Characteristics	Intervention Groups	Outcome	Results
Trials that compare reach strategies with multi-component strategies				
Becker et al., 2008[82] cRCT N_c = 116 N = 1,378 Baseline, 6- and 12-month follow-up Good	Clinical GPs: General practices in semi-rural region of Germany Patients: Patients ages 19 and older presenting symptoms of lower back pain on the day of recruitment	G1: Mailed guideline (Increase clinician reach) G2: Guideline implementation (multi-component, clinicians only) G3: Guideline implementation and motivational counseling directed at patient (multi-component, clinicians and patients)	Overall physical activity (2nd)	Confidence intervals or p values not reported for this comparison
Bishop & Wing, 2006[85] RCT N=462 0-4 weeks, 5-12 weeks, >12 weeks Good	Clinical Patients with acute low back pain symptoms for more than 2 weeks and less than 4 weeks and an accepted claim with the Workers' Compensation Board of British Columbia relating to an injury	G1: Control (not abstracted) G2: Physician only (increase reach) G3: Physician and patient (multi-component)	Guideline-concordant treatment advice and procedures (education, exercise, non-narcaotic medication, spinal manipulation, physical therapy) or guideline-discordant treatment advice and procedures (>4 days bed rest, continued passive therapy, routine use of narcotic analgesics) (P)	Confidence intervals or p values not reported for this comparison
Christakis et al., 2006[100] fRCT N=887 Baseline, up to 365 days after baseline. Fair	Clinical Children under 11 years of age who were patient at a participating clinic and who needed to make a well-child visit during the study period.	G1: Usual care (not abstracted) G2: Parental content Alone (increase reach) G3: Provider notification alone (not abstracted) G4: Parental content and provider notification (multi-component)	Discussion of MyHealthy Child topics (P) Parent mplementation of MyHealthy Child topics (P)	Confidence intervals or p values not reported for this comparison

Table 28. Summary of studies examining dissemination to clinicians and patients for health-related decisions and behavior outcomes (continued)

Author, Year, Design Sample Size Duration Quality Rating	Setting Sample Characteristics	Intervention Groups	Outcome	Results
Feldstein et al., 2006[87] RCT N=327 6 months post intervention Good	Clinical Women ages 50 to 89 who had suffered a fracture in 1999 and had not received bone mineral density measurement or medication for osteoporosis	G1: Usual care (not abstracted) G2: EMR reminder (increase reach for clinicians) G3: EMR reminder and patient reminder (via letter with educational materials (multi-component)	Proportion of study population receiving a pharmacological treatment (P) Proportion of study population receiving bone mineral density measurement within 6 months after the intervention (P) Total calcium intake (2nd) Regular physical activity (2nd) Weekly caloric expenditure (2nd)	G2 = G3 No significant differences between groups For secondary outcomes confidence intervals or p values not reported for these comparisons
Trials that compare ability strategies with multi-component strategies				
Eaton et al., 2011[86] cRCT N_c=30 N=4,239 Good	Clinical Primary care practices in Southeastern New England; other inclusion criteria not reported	G1: 1-hour academic detailing (increase clinician ability) G2: Academic detailing plus a patient education toolkit, a computer kiosk with patient activation software, and a clinician PDA-based decision support tool (multi-component)	Percentage of patients screened for hyperlipidemia treated for their low-density lipoprotein and non–high-density lipoprotein (HDL) cholesterol (P)	G1 = G2 No significant difference between groups

Table 28. Summary of studies examining dissemination to clinicians and patients for health-related decisions and behavior outcomes (continued)

Author, Year, Design Sample Size Duration Quality Rating	Setting Sample Characteristics	Intervention Groups	Outcome	Results
Trials comparing multi-component strategies				
Maxwell et al., 2010[108] cRCT N_c=548 Baseline, 6-month follow-up Fair	Community-based settings Filipino-Americans, ages 50 to 70, who are not current with colorectal cancer screening	G1: Control (not abstracted) G2: Educational session + letter to provider (multi-component) G3: Educational session + letter to provider + FOBT kit (multi-component)	Self-reported CRC screening rates (P)	Confidence intervals or p values not reported for this comparison

Note: In most cases, the labels of the intervention groups represent the labels that study authors had originally used, except if they were merely labeled control or intervention. Information in parentheses represents the strategy that the authors of this review believed best captured the intervention components. When authors used nondescriptive or vague labels for the intervention groups (e.g., Group 1, Intervention A) we used a summary of their description of the group as a label. For a full description of the studies intervention components refer to Appendix E.

Abbreviations: 2nd = secondary outcome; + = Plus; CRC = colorectal cancer; cRCT = cluster randomized controlled trial; EMR, electronic medical record; FOBT, fecal occult blood test; fRCT = factorial randomized controlled trial; Nc, number of clusters; P = primary outcome.

Comparisons of Multi-component Strategies

One cRCT compared two multi-component arms to a control condition in an effort to enhance self-reported CRC screening in a community-based setting.[108] In each arm patients received educational information about CRC and a letter was also sent to the patient's physician. In one of the arms patients also received additional materials in the form of a fecal occult blood test (FOBT) kit. The authors did not make a direct comparison of the multi-component arms to each other, although both were significantly better than the control group.

Detailed Synthesis: Dissemination Strategies for Clinicians and Patients for Clinical Outcomes

Table 29 documents the strength of evidence grading for the dissemination comparison focused on clinicians and patients for clinical outcomes and gives the overall SOE grade. Table 30 describes the single trial in this category.

Table 29. Strength of evidence: Trial examining dissemination to clinicians and patients for clinical outcomes

Strategy	Number of Studies; Subjects; Design/Quality Rating	Risk of Bias	Consistency	Directness	Precision	Results and Strength of Evidence
Reach strategies versus multi-component strategies	1; N_c = 116 cRCT/Good[82]	Low	Unknown (single study)	Direct	Imprecise	Likely no difference between groups; authors did not report confidence intervals or significance tests for comparisons. Insufficient

Abbreviations: cRCT = cluster randomized controlled trial; N_c = number of clusters.

One trial compared a reach strategy (mailed guidelines) with two multi-component strategies that involved providing either education and academic detailing or education, academic detailing, and motivational counseling for patients.[82] This cRCT examined functional capacity among patients with low back pain using 6- and 12-month follow-up assessments. The groups did not differ significantly on functional capacity at either follow-up.

Table 30. Summary of studies examining dissemination to clinicians and patients for clinical outcomes

Author, Year, RefID Design Sample Size Duration Quality Grade	Setting Sample Characteristics	Intervention Groups	Outcome	Results
Studies that compare Reach strategies with Multi-component strategies				
Becker et al., 2008[82] cRCT N_c = 116 N = 1,378 Baseline, 6- and 12-month follow-up Good	Clinical GPs: General practices in semi-rural region of Germany Patients: Patients ages 19 and older presenting symptoms of lower back pain on the day of recruitment	G1: Mailed guideline (Increase clinician reach) G2: Guideline implementation (multi-component, clinicians only) G3: Guideline implementation and motivational counseling directed at patient (multi-component, clinicians and patients)	Functional capacity (Hannover Functional Ability Questionnaire) (P) Quality of Life (2nd)	Confidence intervals or p values not reported for this comparison.

Note: In most cases, the labels of the intervention groups represent the labels that study authors had originally used, except if they were merely labeled control or intervention. Information in parentheses represents the strategy that the authors of this review believed best captured the intervention components. When authors used nondescriptive or vague labels for the intervention groups (e.g., Group 1, Intervention A) we used a summary of their description of the group as a label. For a full description of the studies intervention components refer to Appendix E.

Abbreviations: 2nd = secondary outcome; cRCT = cluster randomized controlled trial; GP, general practitioners; N_c, number of clusters; P = primary outcome.

Results: Key Question 3: Communicating Uncertainty

Introduction

This section presents the results for Key Question (KQ) 3: the effect of alternate ways of communicating uncertainty as it pertains evidence translation. The analytic framework for this question is presented as part of the Introduction and shows the effects of alternate ways of communicating uncertainty on both intermediate and distal outcomes. As we noted in our methods, the best studies to answer this question would be randomized controlled trials (RCTs). However, given the early state of research, we have also included studies with other experimental designs.

In this section, we present our results for four main types of uncertainty: overall strength of evidence as reflected in the wording of various types of clinical recommendations; directness of how evidence is presented; precision with which evidence is provided; and various ways of depicting net benefit of the evidence. We found no eligible studies on risk of bias, consistency, or applicability. Below, we subdivide our results by the specific alternate presentations of uncertainty presented and by outcomes studied to make strength of evidence (SOE) determinations. Tables below describe individual studies and their results and document our SOE grades. Detailed evidence tables for KQ3 are in Appendix F.

Description of Included Studies

Figure 1 in the first results section (for KQ1) depicts the flow of article exclusion and inclusion. After dual review at both the title/abstract and full-text article stage, we retained nine articles,[69,120-127] meeting inclusion criteria and report on eight unique studies about alternative ways of communicating uncertainty.

Of the eight included studies, we graded one as low risk of bias,[120] six as moderate risk of bias,[69,121-126] and one as high risk of bias.[127] We excluded the study with high risk of bias from further analysis.

Below we report on the seven studies with low or moderate risk of bias (i.e., good or fair quality). Of these studies, one was a RCT, four were factorial RCTs, one a non-controlled trial, and one a quasi-experimental study. One reported on the effects of alternate wordings of overall strength of evidence.[122] Three studies reported on various presentations of precision;[121,128] one tested alternative ways of communicating directness[120]; and three investigated different ways of communicating net benefit (with some studies making more than one comparison).[120,123-126] No studies reported on alternate presentations of risk of bias, consistency, or applicability. Three studies reported the effects of alternate non-numeric presentations of uncertainty;[120,122,124] two on alternate numeric presentations;[39,121] one on numeric versus graphical presentations;[39] one on alternate graphical presentations;[39] and 1 on framing.[125]

Interventions were tested in study populations in the United States, Canada, and Switzerland. Sample sizes ranged from 135 participants to 2, 944 participants. Outcomes studied included knowledge, perceived risk, accuracy of perceived risk, appropriate choices regarding care (e.g., selecting medications; obtaining screening), and decision satisfaction.

Key Question 3: The Effects of Communicating Uncertain Health and Health Care Evidence to Patients and Clinicians

Key Points

- **Communicating strength of evidence:** One study found that non-numeric recommendation scheme (e.g. GRADE, NICE, ACCP) was superior to others for all types of recommendations (strong for, weak for, weak against, strong against). Different ways of wording recommendations intended to convey strong recommendations resulted in only moderate guideline-concordant care with no wording superior to others. Weak recommendation wording resulted in higher guideline concordant care with "we suggest" and "we conditionally recommend against" wording performing better than other wording for and against care. (one study; insufficient SOE).
- **Communicating precision:** One study found that presenting numeric risks as point estimates versus 95% confidence intervals (CIs) or changing the format of 95% CIs (numeric versus graphical) did not significantly influence perceived risk of colon cancer (one study; insufficient SOE) or the accuracy of perceived risk of breast cancer recurrence (one study, insufficient evidence) Using clean versus blurry bar graphs to convey information about uncertainty did not produce different levels of perceived risk (one study; low SOE)
- **Communicating directness:** Choice of a cholesterol medication with direct evidence of net benefit was better for patients receiving non-numeric statements that provided either directive advice (i.e. "ask for the drug that…") or factual information about the evidence than for patients receiving usual care. However, medication choices did not differ by type of instruction (one study; low SOE).
- **Communicating net benefit:** Choice of a heartburn medication that was more likely to have net benefit because of its longer history of evidence about safety was better for consumers receiving non-numeric statements that provided either directive advice or factual information rather than usual care, but medication choices did not differ by type of instruction (one study; low SOE). Receiving additional non-numeric information about benefits had little effect on refusals of cancer screening tests, but receiving more non-numeric information on harms significantly increased test refusals and significantly decreased decision satisfaction (all low SOE). Compared with usual care, giving men prostate cancer screening information alone or framed in the context of information about other, more beneficial screening services significantly increased prostate cancer knowledge and the proportion of individuals reporting active involvement in decision-making. However, giving prostate cancer screening information alone increased screening, whereas giving it framed in the broader context of more beneficial services decreased screening (all low SOE).

Tables 31-34 document findings and SOE grades.

Table 31. Strength of evidence: Communicating strength of evidence

Number of Studies; Subjects; Design/Quality	Risk of Bias	Consistency	Directness	Precision	Results and Strength of Evidence
1; 341[122] RCT/ Fair	Moderate	Consistency unknown (single study)	Direct	Imprecise	Relatively few medical residents who received strong non-numeric recommendations for or against care (e.g., "we recommend") reported that they would adhere to guideline-concordant care; there was little difference among guideline concordant care using language from various existing grading schemes (e.g. "we recommend, "clinicians should"). Weak recommendation wording resulted in somewhat higher guideline-concordant care; effects differed by specific wording. For weak recommendations *for* care, "we suggest" performed best. For weak recommendations *against* care, "we conditionally recommend against" performed best. Insufficient

Abbreviations: RCT = randomized controlled trial.

Table 32. Strength of evidence: Communicating precision

Number of Studies; Subjects Design/Quality	Risk of Bias	Consistency	Directness	Precision	Results and Strength of Evidence
Different Numerical Presentations (Point Estimate Versus 95% CI)					
1; 240[69] RCT/ Fair Experiment 1	Moderate	Consistency unknown (single study)	Direct	Imprecise	OUTCOME: Perceived risk When risk is presented alone, no difference in perceived risk with point estimate vs. 95% CI, in either text or graphical format. When risk is presented with comparative information about the risk for an average population, the 95% CI, decreases perceived risk compared with a point estimate in text format and increases perceived risk compared with point estimate in graphical format. Insufficient
1; 143[121] RCT/ Fair	Moderate	Consistency unknown (single study)	Direct	Imprecise	OUTCOME: accuracy of perceived risk No significant differences in the proportion of "gist" or "verbatim" errors (i.e., accuracy of risk perception) when risk estimates were presented as text depictions of the point estimate alone, graphical depictions of the point estimate alone, or as combined text and graphical depiction of the point estimate with accompanying text statement indicating 95% CIs. Insufficient
Numeric Versus Visual Presentations of 95% CI					
1;240[69] RCT/ Fair Experiment 1	Moderate	Consistency unknown (single study)	Direct	Imprecise	No significant difference in perceived risk when 95% CIs were presented in text or in horizontal bar graphs. Insufficient
Different Visual Presentations of 95% CI					
1;135[69] RCT/ Fair Experiment 2	Moderate	Consistency unknown (single study)	Direct	Imprecise	No significant difference in perceived risk when 95% CIs were presented in horizontal bar graphs with solid or hazy borders. Insufficient

Abbreviations: CIs = confidence intervals; RCT = randomized controlled trial.

Table 33. Strength of evidence: Communicating directness

Number of Studies; Subjects Design/Quality	Risk of Bias	Consistency	Directness	Precision	Results and Strength of Evidence
1; 2,944[120] RCT/ Good	Low	Consistency Unknown (single study)	Direct	Precise	Both non-numeric statements that provided directive advice (Ask for a drug that…) and factual information designed to encourage consumers to use cholesterol-lowering drugs with evidence of direct benefit improve appropriate choices compared with usual care. Low

Abbreviations: RCT = randomized controlled trial.

Table 34. Strength of evidence: Communicating net benefit

Number of Studies; Subjects Design/ Quality	Risk of Bias	Consistency	Directness	Precision	Results and Strength of Evidence
Alternate Non-Numeric Statements on Net Benefit					
1; 2,944[120] RCT/ Good	Low	Consistency unknown (single study)	Direct	Precise	Both non-numeric statements that provided directive advice (ask for drug that…) and factual information designed to encourage consumers to use heartburn drugs with a high likelihood of net benefit improve appropriate choices compared with usual care. Low
Varying Amounts of Non-Numeric Benefit and Harm Information					
1; 2,333[123,124] RCT/ Fair	Moderate	Consistency unknown (single study)	Direct	Precise	OUTCOME: decision satisfaction Increasing harms information decreased decision satisfaction. Increasing benefit information had no effect. Low
1; 2, 333[123,124] RCT/ Fair	Moderate	Consistency unknown (single study)	Direct	Precise	OUTCOME: test refusal Increasing harms information increased recipients' refusal of a screening test for an unnamed cancer. Increasing benefit information had no effect. Low
Presenting Framed Net Benefit Information (i.e in Context of Other Services with Different Net Benefit)					
1; 584[125,126] RCT/ Fair	Moderate	Consistency unknown (single study)	Direct	Precise	OUTCOME: knowledge A decision aid with prostate cancer information alone vs. a decision aid with prostate cancer information framed in the context of other, more beneficial screening services increased knowledge of prostate cancer screening. Low
1; 584[125,126] RCT/ Fair	Moderate	Consistency unknown (single study)	Direct	Precise	OUTCOME: screening A decision aid with prostate cancer screening information alone increased screening while a decision aid with prostate cancer screening information framed in the context of other, more beneficial screening services decreased prostate cancer screening. Low
1; 584[125,126] RCT/ Fair	Moderate	Consistency unknown (single study)	Direct	Precise	OUTCOME: involvement in decision-making A decision aid with prostate cancer information alone and one with prostate cancer information framed in the context of other, more beneficial screening services increased active involvement in decision making. However, statistical significance was achieved only in the latter study arm. Low

Abbreviations: RCT = randomized controlled trial

Detailed Synthesis

Communicating Overall Strength of Evidence

One factorial randomized trial examined the effect of alternative non-numeric ways of communicating strength of evidence information on guideline-concordant choices for congestive heart failure or inflammatory bowel disease care (Table 35).[122] Strength of evidence was reflected in the ways that various groups word their recommendations about health care. These mimicked current wording used by the American College of Chest Physicians (e.g., "we recommend," "we suggest"), the National Institute for Health and Clinical Excellence (e.g., "clinicians should", "clinicians might"), or the Grading of Recommendations Assessment, Development, and Evaluation group (e.g., "we recommend," "we conditionally recommend"). In this trial, investigators randomly assigned 341 medical residents to one of three groups with different wording for health care recommendations and subsequently randomized them to receive one strong (i.e., "we recommend") and one weak (i.e., "we suggest") recommendation either "for" or "against" some specific health care option.

The effect of various strong recommendations (e.g. "we recommend" vs. "clinicians should" vs. "we recommend") either for care or against care did not differ much. Relatively few residents who received the strong recommendation either for or against care indicated that they would prescribe care appropriate with the intended recommendation of the guideline (i.e. guideline-concordant care; "for," 8 percent on average across three groups; "against," 47 percent on average across three groups).

Weak recommendation wording yielded somewhat higher guideline-concordant care, but effects differed significantly by specific wording. The choice for guideline-concordant care among residents who received the wording "we suggest" was 31 percentage points higher than among those who received the wording "clinicians should" and 16 percentage points higher than among those who received the wording "we conditionally recommend" ($p=0.003$ for any difference in groups). Further, the choice for guideline-concordant care among residents who received the wording "we conditionally recommend against" was 32 percent higher than in those who received the wording "we recommend against" and 23 percent higher than in those who received the wording "clinicians might" ($p=0.002$ for any difference in groups).

Table 35. Studies of communicating strength of evidence

Author, Year Design Setting Sample Size Quality Grade	Comparator Intervention	Active Interventions	Outcome	Comparator Results	Active Results	Difference Results
Akl, 2012[122] fRCT Residency training programs in United States and Canada Academic health care institutions N = 341 Fair	G3: Strong and weak recommendations for or against guideline-supported behavior (GRADE): • "we recommend" • "we conditionally recommend" • "we conditionally recommend…not" • "we recommend…not"	G1: Strong and weak wording for or against guideline-supported behavior (ACCP): • "we recommend" • "we suggest" • "we suggest...not" "we recommend…not" G2: Strong and weak wording for or against guideline-supported behavior (NICE): • "clinicians should" • "clinicians might" • "clinicians might not" • "clinicians should not"	Appropriate choices of therapy (i.e., consistent with recommendation language), %	Appropriate choices, by wording category: "Strong for" G3: 7% "Weak for "G3: 61% "Weak against" G3: 64% "Strong against" G3: 51%	Appropriate choices, by wording category: "Strong for" G1: 7% G2: 9% "Weak for" G1: 77% G2: 46% "Weak against" G1: 32% G2: 55% "Strong against" G1: 49% G2: 42%	"Strong for" G1 vs. G3: 0%[a] G2 vs. G3: 2%[a] G1 vs. G2: -2%[a] p (overall) = .91 "Weak for" G1 vs. G3: 16%[a] G2 vs. G3: -15%[a] G1 vs. G2: 31%[a] p (overall) = .003 "Weak against" G1 vs. G3: -32%[a] G2 vs. G3: -9%[a] G1 vs. G2: -23%[a] p=.002 "Strong against" G1 vs. G3: -2%[a] G2 vs. G3: -9%[a] G1 vs. G2: 7%[a] p=.60

[a]Calculated by the reviewer

Abbreviations: ACCP = American College of Chest Physicians; fRCT = factorial randomized controlled trial; G = group; GRADE = Grading of Recommendations Assessment, Development, and Evaluation; N = number; NICE = National Institute for Health and Clinical Effectiveness; vs. = versus.

Effects did not differ statistically between the heart failure and inflammatory bowel disease contexts. Based on this small single study, we graded the overall strength of evidence insufficient.

Communicating Precision

Three studies reported on the effect of various ways to communicate the preciseness with which evidence is presented (Table 36). Two reported on the effects of using alternate numeric presentations (95% CIs or point estimates),[69,121] one on the effect of presenting 95% CIs in numeric versus in a visual format,[69] and one on presenting 95% CIs in alternative visual formats.[69]

Intervention: Alternate Numeric Presentations

One factorial randomized trial[69] and one quasi-experiment trial[121] reported the effects of using alternate numeric presentations of precision (CIs versus point estimates) on various outcomes including perceived risk and accuracy of perceived risk. Given their focus on different outcomes, we graded the strength of evidence for these two studies separately.

In the factorial randomized trial, investigators randomly assigned a convenience sample of 240 members of a Web survey panel to receive presentations of either 95% CIs or point estimates describing their hypothetical risk of developing colon cancer.[69] The sample was subsequently randomized to either text or horizontal bar graph presentations. All presentations were done with and without comparative information on the risk of colon cancer in the general population.

In analyses with no comparative risk information, the perceived risk of colon cancer did not differ significantly when risk was presented with 95% CI, versus as a point estimate in either text (+0.3 on 0 to 5 scale, $p>0.05$) or graphical format (-0.5 on 0 to 5 scale, $p>0.05$). However, when risk was presented in the context of comparative risk information, 95% CIs had small, statistically significant effects on perceived risk. This approach reduced perceived risk compared with presentations of point estimates in text format (-0.58 on a 0-5 scale, $p=0.03$) and increased perceived risk compared with presentation of point estimates in the bar graph format (+0.2 on a 0-5 scale, $p=0.03$). Based on this single study focusing on the outcome of perceived risk, we graded the strength of evidence as insufficient.

The results of the quasi-experimental study supplement these findings.[121] It showed that, among a convenience sample of 143 breast cancer patients, a graphical presentation of the point estimate of breast cancer recurrence accompanied by text about the 95% CI, did not affect the accuracy of risk perception compared with a text or graphical presentation of the point estimate alone. Based on this single study focusing on the outcome of accuracy of perceived risk, we graded the strength of evidence as insufficient.

Intervention: Numeric Versus Visual Presentations of Precision

The factorial randomized trial (N=240) mentioned above also examined the effect of numeric versus visual (graphical) presentations of 95% CIs.[69] Perceived risk of colon cancer did not differ significantly when risk was presented in graphical format versus text format (-0.4 on a 0-5 scale, p not significant). Based on this single study, we graded the overall strength of evidence insufficient.

Table 36. Studies of communicating precision

Author, Year Design Setting Sample Size Quality Grade	Comparator Intervention (N)	Active Intervention (N)	Outcome[a]	Comparator Results	Active Results	Difference Results
Han 2011[69] (Experiment 1), fRCT United States, other N = 240 Fair	G1: Point estimate of hypothetical risk of colon cancer in text (''Your chances of developing colon cancer in your lifetime are 9%''). G2: Point estimate of hypothetical risk of colon cancer in horizontal bar graph	G3: Text of range representing confidence intervals for hypothetical risk colon cancer (''Your chances of developing colon cancer in your lifetime are between 5% and 13%''). No point estimate provided. G4: Horizontal bar graph with solid borders depicting range for hypothetical risk of colon cancer. No point estimate provided.	Perceived risk before comparative risk information about general population (range 0-5)	Text G1: 1.7[a] Graph G2: 2.1[a]	Text G3: 2.0[a] Graph G4: 1.6[a]	G3-G1: +0.3, NS G4-G2: -0.5, NS Interaction of uncertainty and representational format: p=0.003
			Perceived risk after comparative risk information about general population (range 0-5)	Text G1: 0.70[a] Graph G2: 0.15[a]	Text G3: 0.12[a] Graph G4: 0.35[a]	G3-G1: -0.58, p=0.03 G4-G2: +0.20, p=0.03 3-way interaction of uncertainty and format and comparative risk information: p<0.001.
Brewer 2012[121] Quasi-Experimental USA, Academic health care institutions, University of North Carolina Breast Clinic N=143 Fair	G1: Point estimate (%) for likelihood of breast cancer recurrence in 10 years with evaluative labels (low, intermediate, or high chance). G2: G1 + risk continuum graphic (i.e. horizontal bar chart) depicting this G6: icon array	G3: G2 + 95% CI, with a verbal translation "chance of recurrence could be as low as 5% or as high as 9% for almost all 95% patients" G5: Standard Oncotype DX report: Test description, cancer recurrence score, a recurrence risk, a graph with recurrence risk + 95% CI, + evaluative labels for risk	Accuracy of risk perception (“gist”), % incorrect/error	Gist errors G1: 13% G2: 6% G6: 16%	Gist errors G3: 5% G5: 17%	Gist errors G1 vs. G3: +8%[b], p value NR G2 vs. G3: +1%, p value NR G6 vs. G3: +11%, p value NR G1 vs. G5: +4%[b]; OR:0.57(0.31 to 1.06),ns G2 vs. G5: +11%[b]; OR: 0.27 (0.12 to 0.58), p<.001 G3 vs. G5: +12%[b] OR:0.23(0.10 to 0.52), p<.001 G5 vs. G6: +1%[b]; OR: 0.79(0.44 to 1.44),ns

Table 36. Studies of communicating precision (continued)

Author, Year Design Setting Sample Size Quality Grade	Comparator Intervention (N)	Active Intervention (N)	Outcome[a]	Comparator Results	Active Results	Difference Results
Brewer 2012[121] (continued)			Accuracy of risk perception ("verbatim") % incorrect/ error	Verbatim errors G1: 8% G2: 7% G6: 21%	Verbatim errors G3: 9% G5: 18%	Verbatim errors G1 vs. G3: +1%[b], p value NR G2 vs. G3: -2%, p value NR G6 vs. G3: -12%, p value NR G1 vs. G5: 1 OR: 0.34(0.18 to 0.64), p<.001 G2 vs. G5: 11%[b]; OR: 0.29(0.16 to 0.52), p<.001 G3 vs. G5: 9%[b]; OR: 0.39(0.21 to 0.74), p<.001 G5 vs. G6: 3%[b]; OR:1.18(0.72 to 1.91), NS

[a] Estimated from figure

[b] Calculated by review team

Abbreviations: CI = confidence interval; DX = diagnosis; fRCT = factorial randomized controlled trial; G = group; N = number; NR = not reported; NS, not significant; OR = odds ratio; USA = United States of America; vs. = versus

Intervention: Alternate Visual Presentations of Precision

One small randomized trial (N=135) examined the effect of alternate horizontal bar graph presentations of 95% CIs: a solid bar graph versus a blurred bar graph that was intended give a better indication of the uncertainty.[69] In this trial, perceived risk of colon cancer did not differ by format (effect size not reported). Based on this single study, we graded the overall strength of evidence insufficient.

Communicating Directness

One large (n=2,934) high-quality (low risk of bias) factorial randomized trial conducted in a nationally representative sample examined the effect of giving non-numeric directive advice versus factual statements to consumers to encourage use of cholesterol-lowering drugs that have proven efficacy in lowering heart attack (the distal and direct outcome) rather than just cholesterol levels (the proximate and indirect outcome) (Table 37).[120] Those in the directive group were told that "Surrogates do not always translate into patient outcomes. Ask for a drug to reduce heart attacks"; those in the nondirective group were told only that "Surrogates do not always translate to patient outcomes." Compared with usual care, both directive advice and factual statements improved appropriate choice of medications (directive: +12 percentage points, 95% CI, 7 to18; nondirective: +12 percentage points, 95% CI, 7 to 18), but medication choices did not differ by type of instruction. Based on this single study with low risk of bias, direct evidence and precise findings, we graded the overall strength of evidence low.

Communicating Net Benefit

Three studies reported on the effects of alternative ways of communicating net benefit (i.e., taking both benefits and harms into account).[120,123-126] One focused on alternate non-numeric ways to promote health care with maximum net benefit, one on what happens to choices and decision satisfaction by varying the amount of non-numeric information about benefits and harms, and one on the effect of presenting net benefit information framed in the context of other, more beneficial health services (Table 38).

Table 37. Study on directness

Author, Year Design Setting Sample Size Quality Grade	Comparator Intervention	Active Intervention	Outcome	Comparator Results	Active Results	Difference Results
Schwartz, 2011[120] fRCT[a] USA, Community-based settings, research panel of ~30,000 households Overall N: 2,944 Good	G1: Control. No explanation about evidence.	G2: Factual statement about evidence ("Surrogates do not always translate into patient outcomes.") G3: Factual statement about the evidence and directive advice about what to do ("Ask for a drug shown to reduce heart attacks.")	Appropriate choice of drug (cholesterol-lowering drug shown to reduce cholesterol and prevent heart attack drug)	Cholesterol-lowering drug G1: 59%	Cholesterol-lowering drug G2: 71% G3: 71%	Absolute difference: G2-G1: +12 %, 95% CI, 7 to 18 G3-G1: +12 %, 95% CI, 7 to 18

Abbreviations: CI = confidence interval; fRCT = factorial randomized controlled trial; G = group; N = number; USA = United States of America

Table 38. Studies about providing information about net benefit (balance of benefits and harms)

Author, Year Design Setting Sample Size Quality Grade	Comparator Intervention	Active Intervention	Outcome	Comparator Results	Active Results	Difference Results
Schwartz, 2011[120] fRCT USA, Community-based settings, research panel of ~30,000 households Overall N: 2944 Good	G1: Control. No explanation about evidence.	G2: Factual statement about the evidence: ("It takes time to establish the safety of new drugs.") G3: Factual statement about the evidence and directive advice about what to do ("Ask for a drug with a longer track record").	Appropriate choice of heartburn drug	Heartburn drug G1: 34%	Heartburn drug G2: 53% G3: 53%	Absolute difference for heartburn drug G2-G1: 19 %, 95% CI, 13 to 24 G3-G1: 19 %, 95% CI, 13 to 24
Perneger 2010[123] and Perneger 2011[123] fRCT Switzerland Community-based settings Overall N: 2,333 Fair	Minimal (i.e. no) risk information (aggregated across groups G1, G2, G3) Minimal (i.e. no) benefit information (aggregated across groups G1, G4, G7)[a]	More than minimal risk information (aggregated across groups G4-G9)[a]: Moderate info: false-positive results A lot of info: false-positive and false-negative results More than minimal benefit information (aggregated across groups G2, G3, G5, G6, G8, G9)[a]: Moderate info: survival benefit A lot of info: survival benefit and reassurance of testing of the screening test.	Decision evaluation (akin to satisfaction; range 0-100), Test refusal (%),	Mean decision satisfaction: Minimal risk, aggregate benefit: 85.9 Minimal benefit, aggregate risk: 81.4	Mean decision satisfaction: Moderate risk, aggregate benefit: 80.4 A lot of risk, aggregate benefit: 81.2 Moderate benefit, aggregate risk: 82.5 A lot of benefit, aggregate risk: 83.6:	Adjusted absolute difference in decision satisfaction: More than minimal vs. minimal risk: -5.1 (-6.6, -3.6) More than minimal vs. minimal benefit: 1.1 (-0.4 to 3.6)

Table 38. Studies about providing information about net benefit (balance of benefits and harms) (continued)

Author, Year Design Setting Sample Size Quality Grade	Comparator Intervention	Active Intervention	Outcome	Comparator Results	Active Results	Difference Results
Perneger 2010[123] and Perneger 2011[123] (continued)				% Test Refusal: Minimal risk, aggregate benefit: 8.8 Minimal benefit, aggregate risk: 16.6	% test refusal: Mod risk, aggregate benefit: 18.9 Lot of risk, aggregate benefit: 21.8 Moderate benefit, aggregate risk: 16.2 Lot of benefit, aggregate risk: 16.3	OR for test refusal (compared with minimal information): Minimal risk info: 1.0 Moderate risk info (FP): 2.5 (1.8 to 3.4) Lot of risk info (FP + FN): 3.0 (2.2 to 4.2) Minimal benefit info: 1.0 Moderate benefit info (survival): 1.0 (0.7 to 1.3) A lot of benefit info (survival and reassurance): 1.0 (0.7 to 1.3)

Table 38. Studies about providing information about net benefit (balance of benefits and harms) (continued)

Author, Year Design Setting Sample Size Quality Grade	Comparator Intervention	Active Intervention	Outcome	Comparator Results	Active Results	Difference Results
McCormack et al., 2011[125] and McCormack 2009[126] Nonrandomized trial Community-based organizations in the US (senior, faith-based, fraternal, fitness, and recreations) Overall N: 584 Fair	G1: Control: usual care for prostate cancer screening	G2: Information on Prostate Cancer Screening Only G3: Information on Prostate Cancer Screening framed in the context of Other More Beneficial Men's Health Services: colorectal cancer screening and cardiovascular screening (includes information on how certain doctors are that men will benefit from screening)	Knowledge at 6 months (range 0-10) Knowledge at 12 months (range 0-10) Active Involvement in Decision Making, % PSA screening at 12 month follow-up, %	Mean knowledge scores (6 months) G1: 3.6	Mean knowledge scores (6 months) G2: 5.1 G3: 4.9	Absolute difference in knowledge (6 months) G3-G1: 1.3[b], p NR G2-G1:1.5[b], p NR G3-G2: 0.2[b], p NR
				Mean knowledge scores (12 months). G1: 3.7	Mean knowledge scores (12 months). G2: 4.5 G3: 4.5	Absolute difference in mean increase from baseline (12 months): G3-G1: +1.5[b] , p<0.001 G2-G1: +0.9[b] , p<0.05
				Decision-making involvement: G1: 75%	Decision-making involvement: G2: 79% G3: 78%	Absolute difference in decision-making involvement: G3-G1: + 3%, p=0.045 G2-G1: + 4%, p=0.064 G3-G2: -1%, p NR
				PSA screening at 12 months G1: 64%	PSA screening at 12 months G2: 71% G3: 61%	Absolute difference in PSA screening (12 months) G3-G1: -3%[b], p NR G2-G1: 7%[b], p NR G3-G2: -10%[b], p NR

[a] Each participant received varying information about the benefits and harms of a screening test for an unnamed cancer. Levels of information on risks and benefits, by group: G1 (control); minimal risk, minimal benefit; G2: minimal risk, moderate benefit; G3: minimal risk, a lot of benefit; G4: moderate risk, minimal benefit info; G5: moderate risk, moderate benefit; G6: moderate risk, a lot of benefit; G7: a lot of risk, minimal benefit; G8: a lot of risk info; moderate benefit; and G9: a lot of risk, a lot of benefit.

[b] Calculated by authors

Abbreviations: FN = false-negative; FP = False-positive; fRCT = factorial randomized controlled trial; G = group; NR = not reported; OR = odds ratio; PSA = prostate-specific antigen.

Intervention: Alternate Non-Numeric Presentations of Net Benefit

One large, good-quality trial mentioned above also examined the effect of giving directive advice or factual statements to consumers to encourage their choice of a heartburn drug more likely to have a net benefit because of its longer evidence of safety.[120] Those in the directive group were told that "It takes time to establish the safety of drugs. Ask for a drug with a longer track record," whereas those in the nondirective group were told only that "It takes time to establish the safety of drugs."

Compared with usual care, both the directive and factual statements improved consumers' appropriate choice of medications (directive: +19 percentage points, 95% CI, 13 to 24; nondirective: +19 percentage points, 95% CI, 13 to 24). The choice of medications did not differ, however, by type of instruction. Based on this single study, we graded the overall strength of evidence low.

Intervention: Varying the Amounts of Non-Numeric Benefit and Harm Information

A large (N=2,333) factorial randomized trial conducted on the general population in one Swiss canton examined the effects of providing varying amounts of non-numeric benefit and harm information on test refusals and decision satisfaction.[123,124] Investigators first randomized participants to receive one of three levels of benefit information about a screening test for an unnamed cancer: no information, information on survival benefit, or information on survival benefit plus information on the relief individuals experience from screening. Participants were subsequently randomized to receive one of three levels of harms information about the same test: no information, information about false-positive results, and information about false-positive and false-negative results.

Although additional benefit information had little effect on test refusals (odds ratio [OR] for full information versus no information: 1.0; 95% CI, 0.7 to 1.3), increasing information on harms significantly increased test refusals (OR for partial vs. no harm information: 2.5; 95% CI, 1.8 to 3.4; OR for full vs. no harm information: 3.0, 95% CI, 2.2 to 4.2). Furthermore, harms (although not benefits) significantly decreased decision satisfaction (-5.1, 95% CI, -6.6 to -3.6 on a scale of 0 to 100). Based on this single study, we graded the overall strength of evidence low.

Intervention: Presenting Net Benefit Framed in the Context of Services with Different Net Benefit

One nonrandomized trial in three communities in North Carolina examined the effect of giving male participants decision aids with information only about prostate cancer screening or about prostate screening framed in the context of other, more beneficial screening tests for men. Outcomes included knowledge, involvement in decision-making, and PSA screening at 12 months.[125,126]

Compared with usual care, both prostate cancer screening information alone and framed in the context of other more beneficial screening services increased prostate cancer knowledge (prostate alone vs. usual care: +0.9 on a 0-10 scale, $p<0.05$; prostate information in context versus usual care: +1.5 on a 0-10 scale, $p<0.001$). Both decision aids also increased the proportion of individuals reporting active involvement in decision-making (prostate information alone versus usual care: +4%, $p=0.064$; prostate information in context versus usual care: +3%, $p=0.045$). However, prostate cancer screening information alone increased screening (+7%, p

NR), whereas prostate cancer screening information framed in the context of the other screening services decreased screening (-3%, p NR). Based on this single study, we graded the overall strength of evidence low.

Discussion

This report presented three separate, but topically related, systematic reviews. The overarching topic involves providing health-related evidence effectively to patients and clinicians. Specifically, we were asked to examine various strategies for communicating and disseminating evidence to these target audiences. As discussed more fully later, communication and dissemination lies on a continuum that ends with effective adoption and implementation. Finally, we were charged with exploring ways to explain uncertainty in evidence about preventive health and health care services.

Some of our findings are pertinent only to one key question (KQ); other findings can inform and provide context for two or all three KQs. In this final section of the report, we first discuss each key question individually in terms of principal findings, limitations and applicability, and future research. In the last part of this chapter, we discuss some issues that cut across all three KQs, noting where important synergies and differences arise, and end with implications.

Key Question 1: Communication Strategies

Key Findings and Strength of Evidence

We found five unique randomized trials (RCTs) examining the effects of alternate communication strategies; these approaches entail tailoring messages to individuals, targeting messages to audience segments, using narratives to convey messages, and using framing to convey messages to various end-users. Overall, the strength of the evidence on communication strategies was poor. Specifically, the trials testing various approaches to framing against either targeting audiences or using narratives (i.e., anecdotal) or statistical evidence did not show long-term differences between groups, and evidence was insufficient for drawing any conclusions. Three trials tested targeting against tailoring messages either alone (n = 2) or in combination with an interpersonal component (n = 1) did not demonstrate that tailoring was not superior to targeting when encouraging screening or changing diet and nutritional behaviors, (all received grades of low strength of evidence).

Findings in Relationship to What is Already Known

Investigators used different strategies to convey evidence (frequently presented as risk and benefit information) to promote informed choices among study participants. The literature suggests that how information is presented can influence health related decisions.[46] On their own, the four strategies we examined have been effective in impacting selected health related outcomes (see Introduction section). We are unable to determine at this time which, if any, of the four strategies are better than any other, or whether or not a particular combination of strategies is better than a single strategy. It may be most prudent to consider using multiple strategies since they are not mutually exclusive.

Nonetheless, our review revealed some intriguing insights about the potential comparative strengths of the four strategies that add to the existing literature. The preponderance of the literature on framing focuses on gain and loss framed messages. It supports that loss-framed messages are more effective than gain frame messages because detection behaviors are perceived as risky in the short term because of their ability to detect disease.[129] Rothman and Salovey[130] concluded that the effect of gain versus loss framing on health behavior was dependent on the function (prevention versus detection) of the behavior.[130] The results of the two trials in our

review that used loss framed prevention messages were consistent with the literature; they both showed loss framed messages were more persuasive, at least in the short-term (6 months). However, both studies used framing in conjunction with another strategy (targeting and narratives).

The included KQ1 trials chiefly involved targeting and tailoring strategies. Investigators hypothesized that tailored interventions would be more effective in promoting screening relative to targeted interventions because they are more personalized. We are aware of at least one study (that did not meet our criteria for evidence) that quantified the contribution of individual tailoring over groups targeting.[131] Two trials directly compared the effectiveness of targeting to tailoring,[77,80] but they produced either nonsignificant or counterintuitive findings. Another trial expected the combination of tailoring and targeting would be more effective than targeting alone,[132] but this was not the case.

The communication strategies we examined are rooted in theory. Framing and targeting may influence beliefs about illness severity and likelihood, which are thought to motivate health behavior change.[133,134] The Health Belief Model[135] supports the notion that providing personalized risk information may increase a person's perceived susceptibility of the disease and motivate their behavior. The belief that a person can undertake a behavior – self-efficacy – is a key variable in numerous health behavior theories[133,135-139] and is frequently an important mediating variable. However, none of the studies examined mediating variables directly, so we could not draw any conclusions about these points.

Narrative communication has emerged a promising approach for cancer prevention and control.[140] Only one trial used some form of narratives in an attempt to improve how effectively the risk and benefit information was conveyed. This study compared using a narrative (anecdotal) approach versus a (non-narrative/statistical) approach, but the investigators did not implement the study design or analyze the data in such a way to uniquely examine each component of the framing versus narratives aspect of the intervention. Instead, they focused on the message framing (gain versus loss). Nonetheless, the results suggest that the loss framed and narrative approach was more persuasive in promoting self-reported likelihood of screening in an immediate post-test. A recent review of the literature found that numeric information appears to improve understanding of risks and benefits relative to non-numeric presentation; presenting both numeric and non-numeric information when possible may be best practice.[61]

In general, few studies set out to examine the comparative effectiveness of these four communication strategies. The strategies are not necessarily substitutes for each other, but are complementary. Using more than one strategy at a time could be synergistic. Thus, one should consider whether looking at comparative effectiveness of two strategies is the most appropriate analytic approach.

Mode of Communication

Most studies delivered the interventions via postal mail or in person. One trial augmented a targeted and tailored intervention by using lay health workers to communicate the information in the intervention to the target audience. In such an approach, communication and dissemination strategies were arguably comingled. Using individuals to help to communicate the messages also did not appear to make any difference in the main outcomes of this trial, even though interpersonal communication is generally viewed as the best approach for communication but also the most expensive.

Advances in technology have made it easier to tailor interventions based on information provided by an individual. Data for tailoring the interventions in the trials were derived from medical record data and from baseline surveys of study participants. Neither of these elements appeared to have any effects on the primary outcomes of interest.

Limitations of the Evidence Base or the Review Process

The evidence base for addressing comparisons of communication strategies of interest was extremely sparse (i.e., only five trials of direct comparisons). Perhaps most telling, four of the five trials addressed breast and cervical cancer screening interventions. The evidence basis for breast cancer screenings has changed in the recent past. As new evidence emerges in the media, such as modifying the target age group or recommended frequency of screening, the result can be confusion among patients[44] and potential interference with the impact of interventions.[44] In addition, one trial had the high baseline screening rates in the study population, so the intervention may have failed to have much of an impact because of a cancer screening ceiling effect.

The fifth trial addressed dietary recommendations. Modifying eating behaviors is notoriously difficult, as is measuring diet changes (e.g., calories from fat; dietary fiber levels). That this one trial (of fair quality) found no statistically significant effects may reflect these challenges.[80]

One major drawback to these trials is the disproportionate use of convenience samples. While randomizing a convenience sample increases the likelihood of equal distribution of characteristics, unmeasured confounding may exist because of selection bias with the sample. All of these trials studies used self-reported data, which can be subject to social desirability bias. For behaviors such as screening or dietary intake, this may pose another limitation.

None of the studies directly examined moderating or mediating variables. Thus, we could not draw any conclusions about their potential role. For example, none of the studies measured the health literacy or numeracy level of the patients which has been shown to be an important factor in comprehending health related information. Similarly, investigators did not control for a variety of potentially confounding variables, such as mode of communication (e.g., use of mailed materials), amount of content (e.g., one-page letter versus multicomponent intervention), or apply modeling techniques that might have clarified the impact (or lack of it) of such factors.

We used rigorous EPC procedures in addressing this KQ that increase the rigor and quality of the review: systematic searches of the literature, appropriate methods for rating quality of studies and grading strength of evidence, rigorous data abstraction techniques, and so forth. We also obtained consultation and input from a wide range of external experts in the field on the development and execution of the review. Despite these strengths, certain tradeoffs were necessary that limited the scope of the review. We emphasize that categorizing the trials proved challenging because the interventions in some cases seemed to be an amalgam of communication techniques. Per our protocol, we did not examine some communication strategies in which some stakeholders might have been interested, such as use of plain language, because we viewed this as a best practice.

Applicability

Applicability can be addressed in terms of PICOTS—i.e., populations, interventions, comparators, outcomes, timeframes for followup, and settings. The generalizability of these findings is limited to the study populations in which they were undertaken. This was chiefly adult women eligible for cancer screening (breast or cervical). No study was relevant to men, to

children or adolescents, or to the elderly. Two trials clearly addressed an ethnic population group (Latinas). The interventions and comparators, as noted above under limitations, were also relatively narrowly conceived. Given the focus on presentation and screening, the outcomes are necessarily quite limited; followup was reasonably lengthy (e.g., 1 year in most cases), and the trials concern chiefly outpatient or community settings. The results of the trials may only be directly applicable to screening and dietary behavior. Thus, we are unable to differentiate findings by socio-economic and clinical characteristics of the study subjects.

Research Gaps

Perhaps a critical step in future research is to reconsider the core communication strategies in light of these findings, theoretical models, research principles and methods, and other indications. The communication strategies were ostensibly evidence based—a crucial factor for selecting interventions or comparators. There was a notable gap in the use of information technology for communicating the information to patients or clinicians. While our analytic framework focused specifically on patients, we recognize that patients and consumers and not necessarily the same, although there is clearly an overlap between these two groups. The findings from this report should be considered in light of this differential and addressed in future research.

Future research should examine the effectiveness of communication strategies in subpopulations including men, not just women. Of particular interest might be investigations that target some health issues or topics that are relevant for both men and women, to help clarify whether different techniques are more (or less) effective depending on sex. Similarly, as the US population ages, research to highlight promising communication approaches for the elderly (including perhaps even persons 85 years of age and older) might be undertaken. Finally, how best to communicate health-related information to children (even quite young children, insofar as they are increasingly conversant with computer and web-based applications) and adolescents will also be important.

As noted above, all five trials for KQ1 focused on preventive health care issues. This might suggest that the body of literature from which these studies emerged is weighted toward preventive health. If so, future research should compare the effectiveness of different communication strategies beyond screening to other content areas.

We applaud that fact that all the included trials except one were quite large (hundreds or thousands of subjects); this feature is important when only relatively small changes or differences between groups are expected. Three of the five had follow-up assessments after the intervention of at least 1 year, which is an important factor in understanding the stability of changes, but two others were relatively short term. We would encourage efforts to measure effects over relatively long periods, especially for trials of communication about disease trajectories, self-management, and treatment to clarify whether the interventions produce lasting impacts

Nevertheless, we note as well that future research should be attentive to methods issues (e.g., study design, analytic and statistical methods). These trials reflect generally accepted methods in the health behavior and education fields, but in some ways they do not report the rigorous standards for health services and policy research. Trials may or may not have implemented these standards in research protocols, but they are not typically reported in the journals that commonly publish these studies. More attention to recruitment and sampling, broader examination of social, psychological and environmental variables that might explain the outcomes, and greater use of

modeling techniques (e.g., multi-level modeling, path analyses) is recommended. Interdisciplinary research teams are likely to strengthen this line of research.

Key Question 2: Dissemination Strategies

Key Findings and Strength of Evidence

We had 37 RCTs, including cluster RCTs, that focused on evidence dissemination to either clinicians or patients (broadly defined) or to both. All used either "single" strategies focusing on increasing the reach, ability, or motivation for the target populations or a multi-component approach with more than one such technique. The aim of these investigations was to enhance health-related decisions or behaviors, clinical outcomes, or knowledge.

Generally, the multi-component intervention approaches had the best strength of evidence (moderate) relative to an intervention relying on a single method (such as a strategy related to increasing reach to clinicians or patients), and particular when focused used for changing clinician behavior. Multi-component intervention approaches for clinicians related to clinical outcomes had low strength of evidence due to inconsistent findings across studies.

Evidence was inconsistent or not statistically significant for all other comparisons for clinicians and patients related to behaviors, clinical outcomes and knowledge resulting in an insufficient strength of evidence judgment for most categories we compared. In addition, the strength of evidence was low or insufficient, often because only a single trial addressed a specific comparison.

We did not find any evidence that any particular single strategy directed at increasing ability or motivation was better than reach strategies. Although here again there were many single studies in these categories that influenced the strength of evidence ratings.

Findings in Relationship to What is Already Known

The findings about the positive impact of multi-component dissemination efforts is consistent with earlier research and prior reviews showing that dissemination strategies that are passive or involve only a single component do not perform as well as more active, multi-component approaches.[35-37] Specifically, we found multicomponent strategies to be more effective than a single strategy for affecting clinicians' behaviors, particularly guideline adherence which was one of the more common outcomes examined. Grimshaw and colleagues conducted the most comprehensive examination of dissemination interventions to date – an overview of systematic reviews – and it also focused on changing provider behavior.[35] In particular, they concluded that multifaceted interventions were more successful when they targeted different barriers to change. Examples of barriers include access to care and the physical environment. Interventions that targeted awareness of these issues and behavior cues to serve as reminders were more promising. However, in subsequent systematic reviews of multifaceted interventions Grimshaw et al.[141] concluded that the effectiveness of multifaceted interventions did not increase incrementally as the number of components increased. As with our included trials they observed that studies did not provide an explicit rationale for why particular interventions, or combinations of components, should work. Our analysis of the multi-component approaches was at a more macro level, and did not account for the exact number or type of components included. A finer grained analysis may have produced results consistent with Grimshaw et al.'s[141] more recent systematic review.

Past research shows the effectiveness of patient-directed interventions decreased with the number of interventions used, whereas the effectiveness of provider-directed interventions increased when up to three interventions were used.[142] The analysis of our included trials did not address the issue. We cannot conclude that specific dissemination tactics to providers or patients or the exact number in a multi-component approach is more effective. However, it is likely that more intense multi-component interventions, i.e., those using more than one additional component, likely were more successful, if they address the determinants of the dissemination problem and provide a clear rationale for why they should work.[35]

Thus, our review is consistent with many of the conclusions offered by Grimshaw and colleagues,[35,141] about strengthening the evidence base for comprehensive multi-component interventions. Our work builds on their work by also suggesting that the finding applies to both primary and secondary outcomes.

None of the trials we examined deemed interventions that involved *both* patients and clinicians as more effective. However, only a limited number of studies examined this combined approach. Prior research indicated that interventions that targeted both professionals and patients were less successful, but interventions that targeted each of these groups individually were both effective.[142] Dissemination studies that examine strategies involving multiple people, such as providers, patients or their family members, and examine the conditions under which dissemination is more successful in changing behavior, would advance this research area. If formative research shows that discussions and communication between patients and providers is an important determinant of the problem then interventions that address their mutual interdependence could work.[143] However, if the determinants of the problem center on separate issues not central to their interactions and communication then targeting each group individually may be effective as well.

As noted above, we categorized studies into the domains of reach, ability, and motivation. One might expect that dissemination strategies that explicitly seek to motivate the target audience to change their behavior would be more effective than those seeking mainly to distribute the information (see Heaney & Israel, 2008[29,144] for review). However, our analysis did not support this conclusion due to insufficient evidence for many of these comparisons due to either a single study in a category, inconsistent findings across studies or non-significant findings. Systematic reviews that address these single strategies would help identify the benefit of these approaches.

Limitations of the Evidence Base or the Review Procedures

Several conceptual and methodological limitations emerged from the literature about dissemination strategies, and some clouded the evidence base about communication as well. First, the significant heterogeneity in how the field references and classifies dissemination strategies continues to confound dissemination and implementation approaches.[6] Authors of other recent reviews have also acknowledged the difficulty in identifying patterns across heterogeneous dissemination studies or data.[28,145]

Not surprisingly, then, the lack of consistency across investigator teams and studies hampered our efforts to classifying a strategy into one of our domain groupings. We had established these domains conceptually at the outset of the work, based on the literature, and with input from a group of technical expects; those domains specified goals of extending the reach, the ability, or the motivation for target populations for single strategies or reflected or multi-component strategies (which might have had more than one such goal).

Second, a related drawback was that many studies, particularly clinically oriented trials, provided little theoretical or conceptual justification for the dissemination strategy employed or a hypothesis for why particular strategy was expected to be more effective than another. With no underlying framework, the investigators, and we, found null findings difficult to interpret and explain. For example, some studies found no effect but cited already high guideline adherence rates by clinicians as the explanation for the lack of impact of the tested dissemination strategy. Understanding why subgroups of clinicians, patients, or other end users, remain resistant to change may benefit from theoretical approaches that draw on social influence, organization theory, and systems science; that is, investigators in the future should (and can) develop stronger rationales underlying the choice of dissemination approaches.[146-152]

Third, studies often confounded the mode of distribution with other variables related to communication. Therefore, we could not tease apart the effect of mode, channel, and content of the intervention on the outcome of interest. As part of our analysis, we mapped each group in every trial to identify the number of components used for that group (single, more than one) and the mode of evidence distribution (e.g., postal mail, face-to-face, electronic, telephone); because the way the evidence was actually conveyed could have been the same or different from mode. For example, postal distribution of evidence could have involved sending a DVD, which *was electronic transmission* of information to the end user.

As noted previously, the overlap between mode of distribution and other elements of interventions blend communication (KQ1) and dissemination (KQ2) together. Evidence dissemination is a specialized form of communication. Future dissemination studies that pay attention to both general communication strategies like those we examined in KQ1 and that address communication variables such as mode, channel and content will advance this area of research and practice.

Fourth, this body of evidence was quite heterogeneous. Across the included studies the health conditions examined included cardiovascular disease, cancer, pain, hypertension, among others. The types of behaviors targeted for change included adherence, smoking cessation, surgery, preventive screening, among others. In addition, we also included various types of clinical outcome and knowledge. Even when investigators studied a similar issue, physical activity or quality of life for example, different measures were used. Moreover, the trials were also diverse with respect to target audience (clinicians, patients, or both). For clinicians, trials may have targeted primary care physicians, residents or nurses. Patients differed on multiple dimensions including the type of behavior change required, the health condition they were managing, and by other background factors such as sociodemographic factors.

To address this heterogeneous and complicated body of work, we classified the trials in broad terms. Nonetheless, this effort still left too few studies in some categories with which to make meaningful conclusions about the relative impact of a particular dissemination strategy. As explained in the results presented in Chapter 4, we could not pursue any quantitative (pooling) analyses, and we summarized results with text, rather than numbers in our summary tables as other investigators have done.[143] (Data supporting our summary can be found in the evidence tables in Appendix E).

Fifth, appropriate follow-up points, including length before final assessment of impact, is critical. Many studies used follow-up periods of six or 12 months. The most appropriate timing for follow-up should be investigated, and the typical follow-up period used in research protocols may not always be the best for testing dissemination hypotheses. Assessments made at an inappropriate time could fail to detect meaningful changes in the outcome of interest. In other

words, if the evaluations were done prematurely, they could have underestimated the true effect size of the intervention. It is important to remember that behavior change is a process, not an event, and the evaluation measures should account for that by monitoring outcomes variables at multiple time points that are specific to the outcome of interest.

Finally, many studies did not consistently compare strategies directly to each other, but instead compared to a usual care or control condition or at times made direct comparisons for only some outcomes. This factor, which limited our ability to draw conclusions about the *comparative* effectiveness of one approach versus another, may be partially an artifact of the conceptual confusion discussed above.

With respect to drawbacks of our review procedures, as noted, we used the same stringent EPC procedures for this KQ on dissemination as noted for KQ1. Per the agreed-upon protocol, we included a broad range of behaviors, clinical conditions, target populations, and dissemination strategies. The resulting heterogeneity in this evidence base actually reflects a commonly encountered attribute of dissemination research.[28,145]

We excluded trials that used strategies that crossed into "implementation research," for example, strategies that related to the point of care, such as clinical reminders or changes in policies and practices that support evidence implementation. Many of our included studies measured outcomes that are closer to those found in implementation research, such as changes in clinical practice or clinical outcomes. These types of outcomes may have underestimated the impact of particular dissemination approaches, because they are conceptually more distal. We recognize that these studies can also inform what we know about the effectiveness of dissemination strategies because the two are integrally linked.

Applicability

As noted above, applicability can be addressed in terms of PICOTS—i.e., populations, interventions, comparators, outcomes, timeframes for follow-up, and settings. The generalizability of these findings is limited to the study populations in which they were undertaken. This was chiefly health care clinicians given the larger number of studies focusing on this population. Studies addressed a variety of target populations: adults with chronic and acute illness, parents, and clinicians of various backgrounds or specialties. No included studies addressed children or adolescents. Few trials addressed an ethnic population group, although two specifically focused on Latinos. The interventions and comparators, as noted above under limitations, were also relatively narrowly conceived. Follow-up was reasonably lengthy, and the trials occurred in multiple setting types.

Research Gaps

To advance dissemination science and practice, future research should address the limitations noted above. Studies should provide a theoretical rationale for strategies used, address and discuss potential confounding, conduct more fine-grained analyses of the determinants of the problem under study, and interpret the findings in greater context given the particular outcome under study. Future research should use a variety of methodologies, beyond RCT and cRCTs to examine dissemination and link dissemination and implementation more closely, and conduct formative dissemination research.[149] These approaches will enhance the external validity of dissemination research.

To improve dissemination, the knowledge base needs to be expanded to incorporate new concepts and methods that will advance dissemination research and practice and link those

efforts with emerging ideas in implementation science (as discussed later in this chapter).[28,151,153] Despite the large number of dissemination studies we identified (certainly relative to communication studies), including previous systematic reviews in the field,[29,34,35] much more needs to be learned about successful dissemination, adoption and implementation of evidence-based practices. More research, using rapid evaluation methodologies,[154,155] would speed evidence generation on this topic and complement the bulk of research in this area that focuses on understanding long-term impacts and stability of those impacts.

Specific research ideas worth pursuing include:

- What types of multi-component approaches, or component combinations, are most effective? Do they affect prevention and treatment behaviors similarly? Does the clinical problem (e.g., cardiovascular disease, cancer, pain) or behavior required to implement the evidence (e.g., guideline adherence, discussions with patients, etc) make a difference?
- What types of communication strategies (e.g., targeting, tailoring, and narratives) can be used in dissemination to clinicians?
- How can formative research be applied to enhance the development of dissemination interventions to enhance effectiveness?
- Can planning models, typically used in public health to develop interventions, also be used to increase the effectiveness of dissemination interventions?
- Do dissemination strategies work equally well in various underrepresented target populations?

The most successful strategy identified in this review was the use of a multi-component dissemination approach for clinicians when trying to change their behaviors. Despite the heterogeneity of studies included in our review, this finding is consistent with other reviews showing the importance of a multi-component approach. As was true in 1995 when Oxman and colleagues[156] published their systematic review on improving professional practice, there is no "magic bullet" for dissemination of evidence. A multi-disciplinary, multi-method approach, but one that harmonizes terminology and methods will advance dissemination research and practice.

Key Question 3: Communication of Uncertainty

Key Findings and Strength of Evidence

In our systematic review on uncertainty, we found seven unique studies that met our inclusion criteria and examined alternate ways to communicate the overall strength of evidence and the precision, directness, and net benefit of evidence. In general, evidence on communicating overall strength of evidence and precision was insufficient. We found no studies examining the effects of alternate ways to communicate information about the risk of bias, consistency, or applicability of evidence.

By contrast, studies examining ways to communicate directness and net benefit demonstrated that several interventions may be helpful in communicating concepts related to uncertainty and improving choices for appropriate care. These approaches include: (1) factual statements and directive advice (i.e. "ask for a drug that…") to choose treatments with direct evidence about benefit (e.g., benefit on ultimate clinical outcomes such as heart attacks, rather than on intermediate outcomes such as cholesterol); (2) factual statements and directive advice (i.e. "ask for a drug that…") to choose treatments with a higher likelihood of net benefit (e.g.,

interventions with longer safety records); and (3) contextual information highlighting the uncertain net benefit of some screening tests relative to other more beneficial tests.

In addition, evidence also suggests that increasing the amount of information presented about the harms of testing affects the choices for testing patients might make. Conversely, increasing the amount of information about benefits has no effect on test choices.

Findings in Relationship to What is Already Known

Our report is the first systematic review about communicating uncertainty inherent in health-related and health care evidence. Although conceptual reviews on communicating uncertainty exist,[53] they focus predominantly on communicating precision and stochastic uncertainty, rather than on the broader concepts related to providing understandable or actionable information about the uncertainty related to evidence. Thus, our review provides the most comprehensive examination of this issue to date.

Specifically, our findings constitute the only rigorous information available about what works (or does not) in conveying the meaning or the data about uncertainty to many different audiences other than those who are technically very sophisticated about these elements of assembling and synthesizing data on clinical or policy topics. Other articles that did not meet our inclusion criteria[66,67,157-161] may be of interest to some working in this area; however, results should be viewed cautiously due to their higher potential for bias and/or lower applicability to targeted end-users for this review.

Limitations in the Evidence Base and Review Procedures

Limitations in the uncertainty literature are evident primarily from the lack of studies that directly test alternate ways to communicate the uncertainty concepts that are relevant to evidence about health and health care. Few studies addressed any type of uncertainty of interest and none examined ways to communicate risk of bias, consistency across studies, or applicability. Further, few alternate communication strategies were tested. Most studies focused on comparing various non-numeric or numeric presentations of uncertainty. Few tested visual presentations or framed presentations and none examined tailored, targeted, or narrative presentations of uncertainty.

To date, the literature has no well-accepted framework defining the important types of uncertainty related to medical evidence or defining the strategies most likely to be effective in communicating about these types of uncertainty. In this review, we have developed a preliminary framework, but this should be vetted by the field.

Further, it is likely that more conceptual work needs to be done to deepen the framework that we developed. For instance, although several studies in our review examined alternate presentations of precision, none examined how presenting precision for both benefits and harms might affect health and medical care. Deepening and centralizing conceptualizations how uncertainty affects use of medical evidence is likely to be quite important.

Applicability

As noted, this evidence has nothing to say about effective ways to communicate uncertainty related to risk of bias assessments, consistency of evidence across studies, or the applicability of evidence itself. Further, evidence regarding communicating uncertainty related to overall strength of evidence and precision was graded as insufficient, making broad application of such evidence premature.

The applicability of evidence for communicating uncertainty about directness and net benefit, which have stronger evidence, is reviewed below. The applicability of evidence regarding using factual statements and directive advice (i.e. "ask for a drug that…") to communicate uncertainty related to directness and overall net benefit is fairly broad given that evidence was derived from large population based studies and effects would not be expected to change across various clinical conditions for prevention and treatment. The applicability of evidence about communicating various numbers of harms and benefits and about presenting net benefit in the context of other more beneficial services should be applied with more caution. Although the evidence is from large population based studies, the specific harms and services communicated may affect outcomes and alter the results when the same interventions are applied in different clinical settings.

Research Gaps

Future studies should address the full spectrum of types of uncertainty related to communicating evidence about health care options. For instance, studies should test a wider variety of possible presentations for strength of evidence; the effects of many current strength of evidence grading schemes have not been tested.

Studies should also target uncertainty for the subdomains comprising overall strength of evidence. Given the lack of any evidence on risk of bias, consistency of evidence, and applicability of evidence, studies in these areas are greatly needed. Additional studies are also needed to test a wider spectrum of ways to communicate precision, directness, and net benefit. Studies of precision should pay particular attention to the effects of presentations of precision for both benefits and harms of interventions, preventive services, and treatment options. Studies of directness should explore how to communicate about indirect patient centered outcomes as well as the more traditional indirect risk factor outcomes. Further, studies of net benefit should pay particular attention to how the number and types of harms used to determine net benefit may affect decisions, to how clinical context affects decisions, and to whether evidence users are using rational or affective decision making to decide about acceptance of therapy.

Future work should also strive for large randomized comparisons of uncertainty presentations to minimize confounding and optimize the precision of findings and for testing of a variety of outcomes ranging from knowledge and accuracy of risk perception to guideline concordant decisions, end-user satisfaction, and trust in the evidence. Testing uncertainty presentations in nationally representative samples or, at the very least, diverse samples will also be critically important given prior work showing differential understanding of information in low- and high-literacy populations.[162]

Evidence in the Broader Context

Many aspects of this review cut across more than one KQ, and some across all three KQs. One aspect of this entire review involves the continuum of generating evidence, synthesizing it, communicating and disseminating findings, and implementing new or revamped practices predicated on those findings. This is a critical conceptual issue. We encountered similarities and differences in the types of comparisons and findings across KQs. Of particular concern are methods issues that reflect recurring gaps and deficiencies in the body of research. Finally, we see certain commonalities in implications for future research and ramifications for clinicians, patients, and other stakeholders and end users.

Below we set findings from our research into the broader context of evidence translation and highlight key cross-cutting issues that might advance the field. We also discuss limitations of our own review that should be considered in interpreting our results.

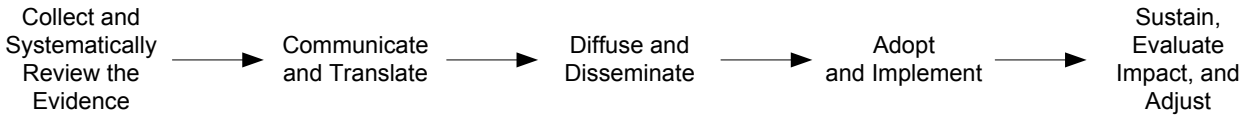
Conceptual Issues

Communication, Dissemination, and Implementation of Evidence

The types of studies and their findings in this synthesis of the literature are best evaluated within a broader context of activities moving from creating interventions, testing them, translating results, and adopting appropriate practices. In health care, this process has been noted as imposing a 17-year period of uptake of proven interventions into routine practice.[163] Experts in the dissemination and implementation science community aim to shorten that period drastically so that health care providers, patients, families, caregivers, communities and healthcare settings are equipped with empirically-supported strategies to integrate scientific knowledge and effective interventions into everyday use. Bernhardt and colleagues[38] lay out the range of basic steps along a spectrum of knowledge creation and use, beginning with the completion of scientific studies to the implementation and maintenance of evidence in practice.[38] Although they focus on Web 2.0 strategies, the key elements are relevant for our more narrow topics and help to illustrate some of the limitations we detected in our evidence base.

In the context of our review, we view the evidence as moving along a continuum beginning with its collection and systematic review followed by communicating and translating it for audiences as needed (see Figure 3). The communication and translation process is often comingled with the diffusion (passive spread) and dissemination (active spread) of the information. Our review included only the second and third phases in the evidence continuum. To be included in our review, studies had to use evidence that met certain criteria for reliability and validity. This requirement increased the quality of the evidence we studied, but it resulted in excluding some studies that may have used effective strategies. Given the scope of our review, we also excluded studies that examined the adoption and implementation phase of the continuum as well as the process of promoting sustainability.

Figure 3. Evidence Continuum in Implementation Science



Each phase in the evidence continuum is unique and essential for evidence to have an impact on public health and health care. Similarly, Lomas argues that "diffusion, dissemination, and implementation are not interchangeable terms: they are phases in a process of increasingly active and more focused intents, with each subsequent phase dependent on the success of its predecessor phase."[8] Researchers, funders, policy makers and the media should keep this full continuum in mind when designing, executing, reviewing, and publishing studies. We cannot be certain that interventions shown to efficacious in a single trial will be effective in real world settings. More emphasis should be placed on conducting comprehensive and longer-term studies that bear all of the phases in mind. Glasgow and colleagues[164] argue that to close the gap between research findings and the application of findings, we also need to balance internal and external study design considerations. With respect to our review, for example, we found that

multicomponent dissemination strategies were generally more effective than single strategies. But, investing in a multicomponent intervention can be resource intensive. Is it reasonable to assume that resources are available for disseminating and implementing multifaceted strategies?

Some trials seemed to conflate communication and dissemination, perhaps not surprisingly given how difficult cleanly defining these concepts can be. Several other trials also seemed to mix or merge dissemination with implementation. This conceptual overlap complicated our analysis in at least two stages: creating meaningful classifications of strategies reported in the literature and examining appropriate relevant outcomes for those strategies.

Better specifying the communication or dissemination process under examination and measuring the most relevant outcomes for that process, and at the relevant follow-up points, will more clearly illuminate the impact of different communication or dissemination strategies. Kreuter and Bernhardt make the same point about the need to reframe the concept of, and the research on, dissemination in particular.[165]

Clearly, as argued also by Green et al.,[149] dissemination is not an end in and of itself[149] and we note that neither is communication. Even if research on these topics bears in mind the full evidence continuum, it needs to be informed by how recipients (or other stakeholders) might view, or use, such materials and knowledge.[29]

Some studies indicated that the investigators had conducted formative work with their target audiences (e.g., patients and consumers, clinicians and other providers, or purchasers and payers); the primary goal was often to identify potential adoptee preferences, values, and needs for evidence. Many research teams did not do this, however. Instead, researchers relied on prior evidence of efficacy for a particular intervention component, especially when it had been "proven" in a similar setting. For example, investigators testing dissemination strategies often used reminders for clinicians about using evidence or following practice guidelines without taking into account whether the potential adoptees had expressed any need for a reminder to do so. Had investigators checked with members of such a target audience, they might have learned that potential adoptees had reasons for ignoring evidence or guidelines, such as perceptions of low credibility or clinical relevance.[166] Armed with such insights, the investigators might have elected to design a different study.

End users of information need to be included in identifying data, findings, and implications to be communicated – i.e., to understand the relevance and usability of such information.[167] Similarly, end users should be involved in developing and testing the evidence product(s) to be communicated, translated, disseminated. Such inclusive practices can help to make the product or information more usable, relevant, and meaningful. Incorporating such stakeholders, specifically including patients or consumers, into future research efforts will require pursuing more participatory research methodologies than has typically been done in the past. This includes involving clinicians in the research process as well given that the format and content of evidence may influence its acceptability and use, and specific content may support shared decision making practices.[153]

Definitions of Concepts and Terms

An important conceptual concern in advancing the science of communicating and disseminating evidence involves definitions of key concepts and terms. As noted in the Introduction to our report, consensus is lacking regarding definitions of key terms pertinent to this review and the research efforts more generally. We saw this lack of consensus across studies especially for definitions of three key terms: dissemination, adoption, and implementation.

For our work, we adopted or adapted definitions used by the US Department of Health and Human Services, particularly the National Institutes of Health, and the health communications literature more generally. Although these definitions differed from each other in minor respects, they were both authoritative and similar enough that we could apply them to frame our inclusion and exclusion criteria, organize our results, and summarize our findings. Greater unity in the field in terms of concepts and terms would be beneficial.

Increasingly, also, investigators in this field are adopting more advanced experimental designs and analytic methods found more commonly today in the health services research and other arenas. Moreover, reflecting this step, research on communication and dissemination is becoming more multidisciplinary. These are very recent advances, however, and our literature went back well beyond recent years.

Use of Theoretical Frameworks and Models

Another important issue for the field is the use of theoretical frameworks. Many studies (but not all) lacked any apparent theoretical or conceptual framework to inform or organize the research questions and focus interventions on essential processes of behavioral and systems change. Although theories and frameworks are distinct concepts, we refer to them collectively, here, as models.[28]

We found it difficult in most studies to determine how investigators used model(s) (if at all) to guide development of their intervention. For example, models did not seem to inform which constructs the researchers chose to frame the intervention, why they chose the constructs they chose, and how they defined those constructions; similarly researchers did not seem to apply models to make predictions about the results of future observations or to interpret study results. Among the trials that did use models, KQ1 studies made greater use of them than did studies for KQ2. KQ1 most frequently used health behavior change models. KQ3 studies were mostly atheoretical.

Existing models had not, of course, been developed to frame research of head-to-head comparisons of communication and dissemination strategies.[166] As a result, investigators sometimes used multiple models, adapting and integrating various components of these models to their particular research questions, settings, intervention design, or intended audiences. These adaptations complicated our efforts to define intervention elements consistently, assure construct validity, and either estimate or accurately measure expected intervention accomplishments and impact, and ultimately to synthesize findings across studies. Adaptations of models can also be barriers to follow-on studies needed to replicate findings, and thus strengthen the body of evidence.

Methodological Considerations Regarding Our Review

In helping readers interpret findings in this report, we must acknowledge the limitations of our own work. Although our review used a well-defined systematic review approach and inputs from key experts, several issues bear mention. Below we highlight issues of which we think readers should be aware.

First, we based our conclusions on findings from RCTs published in peer-reviewed journals for KQ1 and KQ2 and on RCTs plus quasi-experimental and nonrandomized trials for KQ3. We did not consider other sources of information. For instance, the focus on mainly experimental studies necessarily eliminated information from case studies, reports of "exemplar" programs and projects, anecdotal evidence, practice-based (observational) evidence, and consensus expert

opinion. Findings from practice-based initiatives and evaluations of various kinds of programs and projects are frequently of other study designs.

Second, we identified only 53 studies that directly (i.e., head-to-head) compared strategies to communicate and disseminate evidence about evidence. Although we searched multiple medical database using MESH headings and keywords cited by known relevant studies, it is possible that we missed relevant studies that were indexed differently in or in different databases. Given the limited number of direct comparisons, we elected to include studies as long as they had two or more active arms (i.e., study groups). We excluded studies that compared a specific strategy with usual practice alone. However, some studies had multiple arms including a usual practice arm and could have still been included. For example, one KQ1 study had a usual practice arm that was an active comparator (targeting), and many KQ2 studies had a usual practice arm that was an active comparator (e.g., dissemination by mail).

Third, we struggled with what investigators defined as "evidence" to communicate or disseminate. In this review, we excluded studies that disseminated evidence from unknown sources, single studies, or multiple sources of evidence not collected into guidelines, systematic reviews or meta-analyses because we felt they did not reliably represent our intent to study communication and dissemination of evidence-based information. We adopted a policy of including "research-based evidence," which we defined as evidence that various types of authoritative groups (i.e., evidence developers) have assembled, reviewed, and presented and which has been used to make recommendations. Applying this definition allowed us to clearly circumscribe a body of evidence about which we could draw conclusions; however, it may have excluded articles that some fields believe could have been important in answering our key questions.

Fourth, we included a broad range of target populations, interventions (communication or dissemination strategies), outcomes and behaviors, and other variables such as clinical conditions in our work. In short, we cast a "wide net," while attempting to focus clearly on *comparative* effectiveness. As is clear in the results chapters, this heterogeneity in populations, strategies (interventions and comparators), outcomes and methodologies proved to be especially challenging.

Within the RCTs included for KQ1 and KQ2, we found wide variation in the specification of sampling schema, dependent and independent variables, both intermediate and primary outcomes, measurement methods for those outcomes, statistical analyses, and reports of data and findings.

KQ3 studies included more experimental manipulations of different ways of presenting information, especially to lay audiences. This approach is more common in risk communication, the field from which the some of the studies were drawn.

Finally, a mismatch between study design and necessary methodology may partly explain why many of our included studies showed little or no effect of specific intervention strategies. Many of the studies only employed descriptive statistics and did not capitalize on more recent methodological advances (e.g., multi-level modeling) that could have improved their analytic approach. Other studies did not factor in potentially important moderating variables such as self-efficacy and health literacy. As a result of these "negative" findings, we identified few generalizable, efficacious or effective, stand-alone strategies for any of the KQs.

Fifth, a key step in rigorous systematic reviews is assessing the risk of bias (or quality) of individual studies. We adapted our risk of bias assessment from two widely used grading

schemes, although the assessment itself wasn't independently validated. Nonetheless, it covers key domains of selection bias, measurement bias, confounding, power, and reporting issues.

Finally, there is the possibility of publication bias. We did not formally assess publication bias. There were too few studies in any category of strength of evidence to make such assessments meaningful.

Future Research

We have addressed specific research needs for each KQ above. The overall conclusion is that, henceforth, research teams should try to address both the conceptual and the study limitations noted both in sections addressing each key question and here above. This includes: (a) relying more on accepted theoretical constructs and models when designing interventions and studies, (b) conducting some prior needs assessments with target audiences, focusing on audience subgroups with greatest needs, (c) designing robust trials or observational studies, (d) using an array of proven data collection methods that can include, but might go beyond, self-reported attitudes, levels of knowledge, and behaviors; (e) describing and defending choices of intermediate and ultimate outcomes; (f) applying modeling or other advanced statistical and analytic techniques to account for confounders, interactions, and similar complications in data, and addressing temporal aspects of outcomes; and (f) thoroughly describing all aspects of study design and conduct, especially for interventions.

Implications for Patients, Clinicians, and Policymakers

Clinicians and policy makers are in unique positions to use effective communication and dissemination strategies to promote and accelerate the adoption of the evidence base for improving health and healthcare. In many studies, clinicians and policymakers are also members of the intended audience for communication and dissemination strategies designed to alert them to evidence necessitating change in clinical practice and policies guiding it. This review sought to identify effective strategies for communicating and disseminating evidence so that patients and caregivers, practicing clinicians, relevant professional and consumer organizations, purchasers of health care, and others can understand and choose whether to use available evidence as a guide for their clinical practice and policy choices.

Our findings offer some guidance for clinicians and policymakers as to the most effective strategies for communicating and disseminating evidence, but leave many questions unanswered. For example, we found, consistent with other reviews that multi-component strategies addressing a combination of reach, ability, or motivation appear to be more effective than one strategy alone for affecting change in clinicians' behaviors, and particularly clinician guideline adherence (KQ2). Our findings offered us no or insufficient evidence, however to determine the comparative effectiveness of each dissemination strategy within a multi-component strategy. We also found different combinations of strategies with intended audience(s) and setting(s), and few head-to-head comparisons of single strategies, further limiting our ability to recommend a specific strategy or policy for a specific target audience and/or setting.

While clinicians and policymakers may use our findings to guide choice of a specific communication and/or dissemination strategy, they should also carefully consider other factors shown to affect awareness, adoption and use of evidence in various settings and by individuals working in or receiving services in those settings. For example, evidence use by individual clinicians or an organization is dependent on factors such as the definition and source of evidence, the methods used to construct evidence, ways intended audience members use and

retain information, characteristics and expressed needs of the intended audience(s), and organizational as well as individual constraints and enablers specific to various settings. Clinicians and policy makers should gather and use information on these and other factors relevant to their situation or setting as they consider adoption and use of specific communication and dissemination strategies to guide patient-centered care and/or develop and implement systems-level policy.

Patient-centered communication is essential to the provision of patient-centered care and especially so in the face of uncertain evidence about treatment options.[42] We examined the impact of strategies to communicate uncertain evidence including approaches to communicating strength of evidence, precision and directness of risk estimates, and net benefit. Our findings show no consistent impact of these strategies on changing patient outcomes such as knowledge, perceived risk, accuracy of perceived risk, appropriate choices regarding care (e.g., selecting medications; obtaining screening), and decision satisfaction. As clinicians decide how to communicate evidence to their patients, they should consider multi-faceted, multi-directional sharing of information between themselves and their patients that takes into account health literacy, cultural and other factors affecting an individual patient's ability to process and retain information, and use it to make an informed decision. This interaction also hinges on individual clinician knowledge of evidence, organizational policies and standard protocols, and resources available to both the clinician and patient.

The lack of research evidence to inform communication and dissemination of evidence, including uncertain evidence impedes timely clinician, patient and policymaker awareness, uptake and use of evidence to improve the quality of care. Expanding investment in communication and dissemination research is critical to the identification of strategies to accelerate the translation of comparative effectiveness research into community and clinical practice and the direct benefit of patient care.

More research is needed to better understand the current barriers to translating the findings of comparative effectiveness research into community and clinical practice.[168] Ongoing funding for interdisciplinary communication and dissemination sciences research is needed to promote the uptake and use of evidence and ensure quality of care.

Conclusions

In closing, this was the first systematic review that attempted to compare the effectiveness of communication strategies and to look at communicating uncertainty. Finding the appropriate "comparative" studies was challenging. The number of eligible studies was more limited for KQ1 and KQ3, but more substantial for KQ2. Nonetheless, the review provides a helpful foundation in setting the research agenda to address key gaps in the literature.

References

1. Helfand M, Tunis S, Whitlock EP, et al. A CTSA agenda to advance methods for comparative effectiveness research. Clin Transl Sci. 2011 Jun;4(3):188-98. PMID: 21707950.
2. Smith C. The role of health professionals in informing cancer patients: findings from The Teamwork Project (phase one). Health Expectations. 2000 Sep;3(3):217-9. PMID: 11281931.
3. Mitton C, Adair CE, McKenzie E, et al. Knowledge transfer and exchange: review and synthesis of the literature. Milbank Q. 2007 Dec;85(4):729-68. PMID: 18070335.
4. Cabana MD, Rand CS, Powe NR, et al. Why don't physicians follow clinical practice guidelines? A framework for improvement. JAMA. 1999 Oct 20;282(15):1458-65. PMID: 10535437.
5. U.S. Department of Health and Human Services. Health.gov. Washington, DC: Office of the Assistant Secretary for Health, Office of the Secretary, U.S. Department of Health and Human Services. http://www.health.gov/communication/resources/Default.asp. Accessed April 5, 2012.
6. Rabin BA, Brownson RC, Haire-Joshu D, et al. A glossary for dissemination and implementation research in health. J Public Health Manag Pract. 2008 Mar-Apr;14(2):117-23. PMID: 18287916.
7. Brownson RC, Colditz GA, Proctor EK. Dissemination and implementation research in health. New York, NY: Oxford University Press; 2012.
8. Lomas J. Diffusion, dissemination, and implementation: who should do what? Ann N Y Acad Sci. 1993 Dec 31;703:226-35; discussion 35-37. PMID: 8192299.
9. National Institutes of Health. Building the Science of Dissemination and Implementation in the Service of Public Health. 2007 Sep 10-11.
10. McGuire WJ. Personality and attitude change: An information processing theory. In: Greenwald AG, Brock TC, Ostrom TM, eds. Psychological foundations of attitudes. San Diego, CA: Academic Press; 1968:171-96.
11. Wyer RS. The acquisition and use of social knowledge: Basic postulates and representative research Pers Soc Psychol Bull. 1980 December;6:558-73.
12. Petty RE, Cacioppo JT. The Elaboration Likelihood Model of persuasion. In: Berkowitz L, ed. Advances in experimental socialpsychology. Vol. 19. New York: Academic Press; 1986:123-205.
13. Noar SM, Benac CN, Harris MS. Does tailoring matter? Meta-analytic review of tailored print health behavior change interventions. Psychol Bull. 2007 Jul;133(4):673-93. PMID: 17592961.
14. Lustria MLA, Noar SMC, J., Van Stee SK, et al. A meta-analysis of web-delivered, tailored health behavior change interventions. J Health Commun. in press.
15. Slater MD. Choosing segmentation strategies and methods for health communication. In: Maibach E, Parrot EL, eds. Designing health messages. Thousand Oaks: Sage Publications; 1995:186-98.
16. Noar SM, Palmgreen P, Chabot M, et al. A 10-year systematic review of HIV/AIDS mass communication campaigns: Have we made progress? J Health Commun. 2009 Jan-Feb;14(1):15-42. PMID: 19180369.
17. Hinyard LJ, Kreuter MW. Using narrative communication as a tool for health behavior change: A conceptual, theoretical, and empirical overview. Health Educ Behav. 2007 Oct;34(5):777-92. PMID: 17200094.
18. Winterbottom A, Bekker HL, Conner M, et al. Does narrative information bias individual's decision making? A systematic review. Soc Sci Med. 2008 Dec;67(12):2079-88. PMID: 18951673.
19. O'Keefe DJ, Jensen JD. The relative persuasiveness of gain-framed and loss-framed messages for encouraging disease prevention behaviors: a meta-analytic review. J Health Commun. 2007 Oct-Nov;12(7):623-44. PMID: 17934940.

20. Latimer AE, Brawley LR, Bassett RL. A systematic review of three approaches for constructing physical activity messages: What messages work and what improvements are needed? Int J Behav Nutr Phys Act. 2010;7:36. PMID: 20459779.

21. Rimer BK, Kreuter MW. Advancing tailored health communication: A persuasion and message effects perspective. J Commun. 2006;56(s1):S184-S201.

22. Hawkins RP, Kreuter M, Resnicow K, et al. Understanding tailoring in communicating about health. Health Educ Res. 2008 Jun;23(3):454-66. PMID: 18349033.

23. Kreuter MW, Wray RJ. Tailored and targeted health communication: strategies for enhancing information relevance. Am J Health Behav. 2003 Nov-Dec;27 Suppl 3:S227-32. PMID: 14672383.

24. Noar SM, Harrington NG, Aldrich RS. The role of message tailoring in the development of persuasive health communication messages. Communication Yearbook. 2009;33:73-133.

25. Slater MD, Buller DB, Waters E, et al. A test of conversational and testimonial messages versus didactic presentations of nutrition information. J Nutr Educ Behav. 2003 Sep-Oct;35(5):255-9. PMID: 14521825.

26. Wise M, Han JY, Shaw B, et al. Effects of using online narrative and didactic information on healthcare participation for breast cancer patients. Patient Educ Couns. 2008 Mar;70(3):348-56. PMID: 18201859.

27. O'Keefe DJ, Jensen JD. The relative persuasiveness of gain-framed and loss-framed messages for encouraging disease detection behaviors: A meta-analytic review. J Commun. 2009 Jun;59(2):296-316.

28. Tabak RG, Khoong EC, Chambers DA, et al. Bridging research and practice: models for dissemination and implementation research. Am J Prev Med. 2012 Sep;43(3):337-50. PMID: 22898128.

29. Greenhalgh T, Robert G, Macfarlane F, et al. Diffusion of innovations in service organizations: Systematic review and recommendations. Milbank Q. 2004;82(4):581-629. PMID: 15595944.

30. Curry SJ. Organizational interventions to encourage guideline implementation. Chest. 2000 Aug;118(2 Suppl):40S-6S. PMID: 10939998.

31. Ellen ME, Lavis JN, Ouimet M, et al. Determining research knowledge infrastructure for healthcare systems: A qualitative study. Implement Sci. 2011;6(60):60. PMID: 21645401.

32. Green LW, Orleans CT, Ottoson JM, et al. Inferring strategies for disseminating physical activity policies, programs, and practices from the successes of tobacco control. Am J Prev Med. 2006 Oct;31(4 Suppl):S66-81. PMID: 16979471.

33. . Nurse-led cardiac clinics for adults with coronary heart disease. Nurs Stand. 2006 Mar 22-28;20(28):46-50. PMID: 16596861.

34. Bero LA, Grilli R, Grimshaw JM, et al. Closing the gap between research and practice: An overview of systematic reviews of interventions to promote the implementation of research findings. The Cochrane Effective Practice and Organization of Care Review Group. BMJ. 1998 Aug 15;317(7156):465-8. PMID: 9703533.

35. Grimshaw JM, Shirran L, Thomas R, et al. Changing provider behavior: an overview of systematic reviews of interventions. Med Care. 2001 Aug;39(8 Suppl 2):II2-45. PMID: 11583120.

36. Majumdar SR, Soumerai SB. Why most interventions to improve physician prescribing do not seem to work. CMAJ. 2003 Jul 8;169(1):30-1. PMID: 12847036.

37. Sales A, Smith J, Curran G, et al. Models, strategies, and tools. Theory in implementing evidence-based findings into health care practice. J Gen Intern Med. 2006 Feb;21 Suppl 2(Suppl 2):S43-9. PMID: 16637960.

38. Bernhardt JM, Mays D, Kreuter MW. Dissemination 2.0: closing the gap between knowledge and practice with new media and marketing. J Health Commun. 2011;16(Suppl 1):32-44. PMID: 21843094.

39. Han PK, Klein WM, Arora NK. Varieties of uncertainty in health care: a conceptual taxonomy. Med Decis Making. 2011 Nov-Dec;31(6):828-38. PMID: 22067431.

40. Babrow AS, Kasch CR, Ford LA. The many meanings of uncertainty in illness: toward a systematic accounting. Health Commun. 1998;10(1):1-23. PMID: 16370987.

41. Djulbegovic B. Articulating and responding to uncertainties in clinical research. J Med Philos. 2007 Mar-Apr;32(2):79-98. PMID: 17454416.

42. Rowe WD. Understanding Uncertainty. Risk Anal. 1994;14(5):743-50.

43. Mishel MH, Braden CJ. Uncertainty. A mediator between support and adjustment. West J Nurs Res. 1987 Feb;9(1):43-57. PMID: 3647707.

44. Squiers LB, Holden DJ, Dolina SE, et al. The public's response to the U.S. Preventive Services Task Force's 2009 recommendations on mammography screening. Am J Prev Med. 2011 May;40(5):497-504. PMID: 21496748.

45. McGettigan P, Sly K, O'Connell D, et al. The effects of information framing on the practices of physicians. J Gen Intern Med. 1999 Oct;14(10):633-42. PMID: 10571710.

46. Edwards A, Elwyn G, Covey J, et al. Presenting risk information—a review of the effects of "framing" and other manipulations on patient outcomes. Journal of Health Communication. 2001 Jan-Mar;6(1):61-82. PMID: 11317424.

47. Moxey A, O'Connell D, McGettigan P, et al. Describing treatment effects to patients. J Gen Intern Med. 2003 Nov;18(11):948-59. PMID: 14687282.

48. Covey J. A meta-analysis of the effects of presenting treatment benefits in different formats. Medical decision making : an international journal of the Society for Medical Decision Making. 2007 Sep-Oct;27(5):638-54. PMID: 17873250.

49. Visschers VH, Meertens RM, Passchier WW, et al. Probability information in risk communication: a review of the research literature. Risk Anal. 2009 Feb;29(2):267-87. PMID: 19000070.

50. Cuite CL, Weinstein ND, Emmons K, et al. A test of numeric formats for communicating risk probabilities. Med Decis Making. 2008 May-Jun;28(3):377-84. PMID: 18480036.

51. Woloshin S, Schwartz LM. Communicating data about the benefits and harms of treatment: a randomized trial. Ann Intern Med. 2011 Jul 19;155(2):87-96. PMID: 21768582.

52. Garcia-Retamero R, Galesic M. Communicating treatment risk reduction to people with low numeracy skills: a cross-cultural comparison. Am J Public Health. 2009 Dec;99(12):2196-202. PMID: 19833983.

53. Politi MC, Han PK, Col NF. Communicating the uncertainty of harms and benefits of medical interventions. Med Decis Making. 2007 Sep-Oct;27(5):681-95. PMID: 17873256.

54. Mishel MH, Braden CJ. Finding meaning: antecedents of uncertainty in illness. Nurs Res. 1988 Mar-Apr;37(2):98-103, 27. PMID: 3347527.

55. Tannert C, Elvers HD, Jandrig B. The ethics of uncertainty. In the light of possible dangers, research becomes a moral duty. EMBO Rep. 2007 Oct;8(10):892-6. PMID: 17906667.

56. Owens DK, Lohr KN, Atkins D, et al. Grading the Strength of a Body of Evidence When Comparing Medical Interventions. Methods Guide for Effectiveness and Comparative Effectiveness Reviews. Rockville, MD: Agency for Healthcare Research and Quality; 2008.

57. Harris RP, Helfand M, Woolf SH, et al. Current methods of the US Preventive Services Task Force: a review of the process. Am J Prev Med. 2001 Apr;20(3 Suppl):21-35. PMID: 11306229.

58. Sawaya GF, Guirguis-Blake J, LeFevre M, et al. Update on the methods of the U.S. Preventive Services Task Force: estimating certainty and magnitude of net benefit. Ann Intern Med. 2007 Dec 18;147(12):871-5. PMID: 18087058.

59. Babrow AS, Kline KN. From "reducing" to "coping with" uncertainty: reconceptualizing the central challenge in breast self-exams. Soc Sci Med. 2000 Dec;51(12):1805-16.

60. Glenton C, Santesso N, Rosenbaum S, et al. Presenting the results of Cochrane Systematic Reviews to a consumer audience: a qualitative study. Med Decis Making. 2010 Sep-Oct;30(5):566-77. PMID: 20643912.

61. West SL, Squiers LB, McCormack L, et al. Communicating Quantitative Risks and Benefits in Promotional Prescription Drug Labeling or Print Advertising.

62. Agency for Healthcare Research and Quality. Effective Health Care: What Is the Effective Health Care Program. Rockville, MD: US Department of Health and Human Services; 2012. http://www.effectivehealthcare.ahrq.gov/index.cfm/what-is-the-effective-health-care-program1/. Accessed March 29.

63. Institute of Medicine. Initial priorities for comparative effectiveness research. Washington, DC: The National Academies Press; 2000.

64. Agency for Healthcare Research and Quality. Methods Guide for Effectiveness and Comparative Effectiveness Reviews. Agency for Healthcare Research and Quality AHRQ Publication No. 10(11)-EHC063-EF. Rockville, MD: February 2011. Chapters available at: www.effectivehealthcare.ahrq.gov.

65. Kick EL, McKinney LA, McDonald S, et al. A Multiple-Network Analysis of the World System of Nations, 1995-1999. In: Scott J, Carrington P, eds. Sage Handbook of Social Network Analysis. Thousand Oaks, CA: Sage Publications; 2011:311-27.

66. Kuhn KM, Budescu DV. The relative importance of probabilities, outcomes, and vagueness in hazard risk decisions. Organ Behav Hum Decis Process. 1996;68(3):301-17.

67. Ibrekk H, Morgan MG. Graphical communication of uncertain quantities to nontechnical people. Risk Anal. 1987;7(4):519-29.

68. Steginga SK, Occhipinti S. The application of the heuristic-systematic processing model to treatment decision making about prostate cancer. Med Decis Making. 2004 Nov-Dec;24(6):573-83. PMID: 15534339.

69. Han PKJ, Klein WMP, Lehman T, et al. Communication of uncertainty regarding individualized cancer risk estimates: effects and influential factors. Med Decis Making. 2011 Mar-Apr;31(2):354-66. PMID: 20671211.

70. Moher D, Liberati A, Tetzlaff J, et al. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. J Clin Epidemiol. 2009 Oct;62(10):1006-12. PMID: 19631508.

71. Viswanathan M, Berkman ND. Development of the RTI item bank on risk of bias and precision of observational studies. J Clin Epidemiol. 2012 Feb;65(2):163-78. PMID: 21959223.

72. Higgins JPT, Altman DG. Assessing risk of bias in included studies. Cochrane Handbook for Systematic Reviews of Interventions. Hoboken, NJ: John Wiley & Sons; 2008:187-241.

73. Owens DK, Lohr KN, Atkins D, et al. AHRQ series paper 5: Grading the strength of a body of evidence when comparing medical interventions--Agency for Healthcare Research and Quality and the Effective Health-Care program. J Clin Epidemiol. 2010 May;63(5):513-23. PMID: 19595577.

74. Atkins D, Chang SM, Gartlehner G, et al. Assessing applicability when comparing medical interventions: AHRQ and the Effective Health Care Program. J Clin Epidemiol. 2011 Nov;64(11):1198-207. PMID: 21463926.

75. Cox D, Cox AD. Communicating the consequences of early detection: The role of evidence and framing. J Market. 2001;65(3):91-103.

76. Schneider TR, Salovey P, Apanovitch AM, et al. The effects of message framing and ethnic targeting on mammography use among low-income women. Health Psychol. 2001;20(4):256-66. PMID: 11515737.

77. Jibaja-Weiss ML, Volk RJ, Kingery P, et al. Tailored messages for breast and cervical cancer screening of low-income and minority women using medical records data. Patient Educ Couns. 2003 Jun;50(2):123-32. PMID: 12781927.

78. Vernon SW, del Junco DJ, Tiro JA, et al. Promoting regular mammography screening II. Results from a randomized controlled trial in US women veterans. J Natl Cancer Inst. 2008 Mar 5;100(5):347-58. PMID: 18314474.

79. del Junco DJ, Vernon SW, Coan SP, et al. Promoting regular mammography screening I. A systematic assessment of validity in a randomized trial. J Natl Cancer Inst. 2008 Mar 5;100(5):333-46. PMID: 18314473.

80. Elder JP, Ayala GX, Campbell NR, et al. Long-term effects of a communication intervention for Spanish-dominant Latinas. Am J Prev Med. 2006 Aug;31(2):159-66. PMID: 16829333.

81. Bahrami M, Deery C, Clarkson JE, et al. Effectiveness of strategies to disseminate and implement clinical guidelines for the management of impacted and unerupted third molars in primary dental care, a cluster randomised controlled trial. Br Dent J. 2004 Dec 11;197(11):691-6; discussion 88. PMID: 15592551.

82. Becker A, Leonhardt C, Kochen MM, et al. Effects of two guideline implementation strategies on patient outcomes in primary care: a cluster randomized controlled trial. Spine (Phila Pa 1976). 2008(5):473-80. PMID: 18317189.

83. Bekkering GE, Hendriks HJ, van Tulder MW, et al. Effect on the process of care of an active strategy to implement clinical guidelines on physiotherapy for low back pain: a cluster randomised controlled trial. Qual Saf Health Care. 2005(2):107-12. PMID: 15805455.

84. Bekkering GE, van Tulder MW, Hendriks EJ, et al. Implementation of clinical guidelines on physical therapy for patients with low back pain: randomized trial comparing patient outcomes after a standard and active implementation strategy. Phys Ther. 2005(6):544-55. PMID: 15921475.

85. Bishop PB, Wing PC. Knowledge transfer in family physicians managing patients with acute low back pain: a prospective randomized control trial. Spine J. 2006 May-Jun;6(3):282-8. PMID: 16651222.

86. Eaton CB, Parker DR, Borkan J, et al. Translating cholesterol guidelines into primary care practice: a multimodal cluster randomized trial. Ann Fam Med. 2011 Nov-Dec;9(6):528-37. PMID: 22084264.

87. Feldstein A, Elmer PJ, Smith DH, et al. Electronic medical record reminder improves osteoporosis management after a fracture: a randomized, controlled trial. J Am Geriatr Soc. 2006 Mar;54(3):450-7. PMID: 16551312.

88. Jousimaa J, Mäkelä M, Kunnamo I, et al. Primary care guidelines on consultation practices: the effectiveness of computerized versus paper-based versions. A cluster randomized controlled trial among newly qualified primary care physicians. Int J Technol Assess Health Care. 2002 Sum(3):586-96. PMID: 12391951.

89. Kennedy ADM, Sculpher MJ, Coulter A, et al. A multicentre randomised controlled trial assessing the costs and benefits of using structured information and analysis of women's preferences in the management of menorrhagia. Health Technol Assess. 2003;7(8):1-76. PMID: 12633525.

90. Lien LF, Brown AJ, Ard JD, et al. Effects of PREMIER lifestyle modifications on participants with and without the metabolic syndrome. Hypertension. 2007 Oct;50(4):609-16. PMID: 17698724.

91. Young DR, Vollmer WM, King AC, et al. Can individuals meet multiple physical activity and dietary behavior goals? Am J Health Behav. 2009 May-Jun;33(3):277-86. PMID: 19063649.

92. Svetkey LP, Harsha DW, Vollmer WM, et al. Premier: A clinical trial of comprehensive lifestyle modification for blood pressure control: rationale, design and baseline characteristics. Ann Epidemiol. 2003 Jul;13(6):462-71. PMID: 12875806.

93. Rebbeck T, Maher CG, Refshauge KM. Evaluating two implementation strategies for whiplash guidelines in physiotherapy: a cluster randomised trial. Aust J Physiother. 2006;52(3):165-74. PMID: 16942451.

94. Simon SR, Majumdar SR, Prosser LA, et al. Group versus individual academic detailing to improve the use of antihypertensive medications in primary care: a cluster-randomized controlled trial. Am J Med. 2005 May;118(5):521-8. PMID: 15866255.

95. Wolters R, Wensing M, Klomp M, et al. Effects of distance learning on clinical management of LUTS in primary care: a randomised trial. Patient Educ Couns. 2005 Nov;59(2):212-8. PMID: 16257627.

96. Wright FC, Gagliardi AR, Law CH, et al. A randomized controlled trial to improve lymph node assessment in stage II colon cancer. Arch Surg. 2008 Nov;143(11):1050-5; discussion 5. PMID: 19015462.

97. Beaulieu MD, Brophy J, Jacques A, et al. Drug treatment of stable angina pectoris and mass dissemination of therapeutic guidelines: a randomized controlled trial. QJM. 2004 Jan;97(1):21-31. PMID: 14702508.

98. Campbell MK, James A, Hudson MA, et al. Improving multiple behaviors for colorectal cancer prevention among African American church members. Health Psychol. 2004;23(5):492-502. PMID: 15367069.

99. Carney PA, Harwood BG, Greene MA, et al. Impact of a telephone counseling intervention on transitions in stage of change and adherence to interval mammography screening (United States). Cancer Causes Control. 2005 Sep;16(7):799-807. PMID: 16132790.

100. Christakis DA, Zimmerman FJ, Rivara FP, et al. Improving pediatric prevention via the internet: a randomized, controlled trial. Pediatrics. 2006 Sep;118(3):1157-66. PMID: 16951011.

101. Davis J, Roberts R, Davidson DL, et al. Implementation strategies for a Scottish national epilepsy guideline in primary care: results of the Tayside Implementation of Guidelines in Epilepsy Randomized (TIGER) trial. Epilepsia. 2004 Jan;45(1):28-34. PMID: 14692904.

102. Gattellari M, Ward JE. A community-based randomised controlled trial of three different educational resources for men about prostate cancer screening. Patient Educ Couns. 2005;57(2):162-82. PMID: 15911190.

103. Hagmolen of ten Have W, van den Berg NJ, van der Palen J, et al. Implementation of an asthma guideline for the management of childhood asthma in general practice: a randomised controlled trial. Prim Care Respir J. 2008 Jun;17(2):90-6. PMID: 18274692.

104. Jain MK, Heyland D, Dhaliwal R, et al. Dissemination of the Canadian clinical practice guidelines for nutrition support: results of a cluster randomized controlled trial. Crit Care Med. 2006 Sep;34(9):2362-9. PMID: 16850001.

105. Junghans C, Feder G, Timmis AD, et al. Effect of patient-specific ratings vs conventional guidelines on investigation decisions in angina: Appropriateness of Referral and Investigation in Angina (ARIA) Trial. Arch Intern Med. 2007 Jan 22;167(2):195-202. PMID: 17242322.

106. Laprise R, Thivierge R, Gosselin G, et al. Improved cardiovascular prevention using best CME practices: a randomized trial. J Contin Educ Health Prof. 2009 Winter;29(1):16-31. PMID: 19288563.

107. Marcus BH, Napolitano MA, King AC, et al. Telephone versus print delivery of an individualized motivationally tailored physical activity intervention: Project STRIDE. Health Psychol. 2007 Jul;26(4):401-9. PMID: 17605559.

108. Maxwell AE, Bastani R, Danao LL, et al. Results of a community-based randomized trial to increase colorectal cancer screening among Filipino Americans. Am J Public Health. 2010 Nov;100(11):2228-34. PMID: 20864724.

109. Murtaugh CM, Pezzin LE, McDonald MV, et al. Just-in-time evidence-based e-mail "reminders" in home health care: impact on nurse practices. Health Serv Res. 2005 Jun;40(3):849-64. PMID: 15960694.

110. Paradis HA, Conn KM, Gewirtz JR, et al. Innovative delivery of newborn anticipatory guidance: a randomized, controlled trial incorporating media-based learning into primary care. Acad Pediatr. 2011 Jan-Feb;11(1):27-33. PMID: 21272821.

111. Partin MR, Nelson D, Radosevich D, et al. Randomized trial examining the effect of two prostate cancer screening educational interventions on patient knowledge, preferences, and behaviors. J Gen Intern Med. 2004 Aug;19(8):835-42. PMID: 15242468.

112. Rahme E, Choquette D, Beaulieu M, et al. Impact of a general practitioner educational intervention on osteoarthritis treatment in an elderly population. Am J Med. 2005 Nov;118(11):1262-70. PMID: 16271911.

113. Rimer BK, Halabi S, Sugg Skinner C, et al. The short-term impact of tailored mammography decision-making interventions. Patient Educ Couns. 2001 Jun;43(3):269-85. PMID: 11384825.

114. Soler N, Ballester E, Martin A, et al. Changes in management of chronic obstructive pulmonary disease (COPD) in primary care: EMMEPOC study. Respir Med. 2010 Jan;104(1):67-75. PMID: 20122630.

115. Sullivan MD, Gaster B, Russo J, et al. Randomized trial of web-based training about opioid therapy for chronic pain. Clin J Pain. 2010 Jul;26(6):512-7. PMID: 20551726.

116. Watson MC, Bond CM, Grimshaw JM, et al. Educational strategies to promote evidence-based community pharmacy practice: a cluster randomized controlled trial (RCT). Fam Pract. 2002 Oct;19(5):529-36. PMID: 12356708.

117. Banait G, Sibbald B, Thompson D, et al. Modifying dyspepsia management in primary care: a cluster randomised controlled trial of educational outreach compared with passive guideline dissemination. Br J Gen Pract. 2003 Feb;53(487):94-100. PMID: 12817353.

118. King AC, Friedman R, Marcus B, et al. Ongoing physical activity advice by humans versus computers: The Community Health Advice by Telephone (CHAT) trial. Health Psychol. 2007 Nov;26(6):718-27. PMID: 18020844.

119. Wetter DW, Mazas C, Daza P, et al. Reaching and treating Spanish-speaking smokers through the National Cancer Institute's Cancer Information Service. A randomized controlled trial. Cancer. 2007 Jan 15;109(2 Suppl):406-13. PMID: 17149758.

120. Schwartz LM, Woloshin S. Communicating uncertainties about prescription drugs to the public: a national randomized trial. Arch Intern Med. 2011 Sep 12;171(16):1463-8. PMID: 21911629.

121. Brewer NT, Richman AR, DeFrank JT, et al. Improving communication of breast cancer recurrence risk. Breast Cancer Res Treat. 2012 Jun;133(2):553-61. PMID: 21964579.

122. Akl EA, Guyatt GH, Irani J, et al. "Might" or "suggest"? No wording approach was clearly superior in conveying the strength of recommendation. J Clin Epidemiol. 2012 Mar;65(3):268-75. PMID: 22075112.

123. Perneger TV, Schiesari L, Cullati S, et al. Does information about risks and benefits improve the decision-making process in cancer screening - randomized study. Cancer Epidemiol. 2011 Dec;35(6):574-9. PMID: 21622043.

124. Perneger TV, Cullati S, Schiesari L, et al. Impact of information about risks and benefits of cancer screening on intended participation. Eur J Cancer. 2010 Aug;46(12):2267-74. PMID: 20466537.

125. McCormack L, Treiman K, Bann C, et al. Translating medical evidence to promote informed health care decisions. Health Serv Res. 2011 Aug;46(4):1200-23. PMID: 21352225.

126. McCormack LA, Bann CM, Williams-Piehota P, et al. Communication message strategies for increasing knowledge about prostate cancer screening. J Cancer Educ. 2009;24(3):238-43. PMID: 19526414.

127. Johnson BB, Slovic P. Presenting uncertainty in health risk assessment: Initial studies of its effects on risk perception and trust. Risk Anal. 1995 Aug;15(4):485-94. PMID: 7480948.

128. Kylat Ranjit I, Ohlsson A. Recombinant human activated protein C for severe sepsis in neonates. Cochrane Database of Systematic Reviews: John Wiley & Sons, Ltd; 2012.

129. Rothman A, Salovey P, Antone C, et al. The influence of message framing on intentions to perform health behaviors. J Exp Soc Psychol. 1993;29:408-33.

130. Rothman AJ, Salovey P. Shaping perceptions to motivate healthy behavior: the role of message framing. Psychol Bull. 1997 Jan;121(1):3-19. PMID: 9000890.

131. Ryan GL, Skinner CS, Farrell D, et al. Examining the boundaries of tailoring: The utility of tailoring versus targeting mammography interventions for two distinct populations. Health Educ Res. 2001 Oct;16(5):555-66. PMID: 11675803.

132. Vernon SW, Bartholomew LK, McQueen A, et al. A randomized controlled trial of a tailored interactive computer-delivered intervention to promote colorectal cancer screening: sometimes more is just the same. Ann Behav Med. 2011 Jun;41(3):284-99. PMID: 21271365.

133. Rogers RW. Cognitive and physiological processes in fear appeals and attitude change: A revised theory of protection motivation. In: Cacioppo JT, Petty RE, eds. Social psychophysiology. New York: Guilford; 1983:153-76.

134. Rosenstock IM. Historical origins of the health belief model. Health Educ Monogr. 1974;2:328-35.

135. Rosenstock IM, Strecher VJ, Becker MH. Social learning theory and the Health Belief Model. Health Educ Q. 1988 Summer;15(2):175-83. PMID: 3378902.

136. Bandura A. Self-efficacy: The exercise of control. New York: Freeman; 1997.

137. Fishbein M, Ajzen I. Predicting and changing behavior: The reasoned action approach. New York: Psychology Press; 2010.

138. Lazarus RS, Cohen JB. Environmental stress. In: Altman I, Wohlwill JF, eds. Human behavior and environment. Vol. 2. New York: Plenum; 1977:89-127.

139. Prochaska JO. Systems of psychotherapy: A transactional analysis. 2nd ed., Pacific Grove, CA: Brooks-Cole; 1984.

140. Kreuter MW, Green MC, Cappella JN, et al. Narrative communication in cancer prevention and control: a framework to guide research and application. Ann Behav Med. 2007 Jun;33(3):221-35. PMID: 17600449.

141. Grimshaw G, Stanton A. Tobacco cessation interventions for young people. Cochrane Database of Systematic Reviews: John Wiley & Sons, Ltd; 2006.

142. Lancaster T, Silagy C, Gray S, et al. The effectiveness of training health professional to provide smoking cessation interventions: systematic review of randomized controlled trials (Cochrane Review). Oxford: Update Software; 1998.

143. Lewis MA, DeVellis B, Sleath B. Interpersonal communication and social influence. In: Glanz K, Lewis FM, Rimer BK, eds. Health behavior and health education: Theory, research, and practice. 3rd ed. San Francisco: Jossey Bass; 2002:363–402.

144. Heaney CA, Israel BA. Social networks and social support. In: Glanz K, Rimer BK, Viswanath K, eds. Health behavior and health education: Theory, research, and practice. 4th ed. San Francisco, CA: Jossey-Bass; 2008.

145. Powell BJ, McMillen JC, Proctor EK, et al. A compilation of strategies for implementing clinical innovations in health and mental health. Med Care Res Rev. 2012 Apr;69(2):123-57. PMID: 22203646.

146. Bartholomew LK, Parcel GS, Kok G, et al. Intervention Mapping: Designing theory-and evidence-based health promotion programs. Mountain View, CA: Mayfield Publishing Company; 2001.

147. Frambach RT, Schillewaert N. Organizational innovation adoption: A multi-level framework of determinants and opportunities for future research. J Bus Res. 2002;55:163-76.

148. Weiner BJ, Lewis MA, Clauser SB, et al. In search of synergy: strategies for combining interventions at multiple levels. J Natl Cancer Inst. 2012;44:34-41.

149. Green LW, Ottoson JM, Garcia C, et al. Diffusion theory and knowledge dissemination, utilization, and integration in public health. Annu Rev Public Health. 2009;30:151-74. PMID: 19705558.

150. Mendel P, Meredith LS, Schoenbaum M, et al. Interventions in organizational and community context: A framework for building evidence on dissemination and implementation in health services research. Adm Policy Ment Health. 2008;35:21-37.

151. Wandersman A, Duffy J, Flaspohler P, et al. Bridging the gap between prevention research and practice: the interactive systems framework for dissemination and implementation. Am J Community Psychol. 2008 Jun;41(3-4):171-81. PMID: 18302018.

152. Dearing JW, Maibach EW, Buller DB. A convergent diffusion and social marketing approach for disseminating proven approaches to physical activity promotion. Am J Prev Med. 2006 Oct;31(4 Suppl):S11-23. PMID: 16979466.

153. Gagliardi AR, Brouwers MC, Palda VA, et al. How can we improve guideline use? A conceptual framework of implementability. Implement Sci. 2011;6:26. PMID: 21426574.

154. McMullen CK, Ash JS, Sittig DF, et al. Rapid assessment of clinical information systems in the healthcare setting: an efficient method for time-pressed evaluation. Methods Inf Med. 2011;50(4):299-307. PMID: 21170469.

155. McNall M, Foster-Fishman PG. Methods of rapid evaluation, assessment, and appraisal. Am J Eval. 2007 Jun;28(2):151-68.

156. Oxman AD, Thomson MA, Davis DA, et al. No magic bullets - a systematic review of 102 trials of interventions to improve professional practice. Can Med Assoc J. 1995 Nov 15;153(10):1423-31. PMID: 7585368.

157. Akl EA, Maroun N, Guyatt G, et al. Symbols were superior to numbers for presenting strength of recommendations to health care consumers: a randomized trial. J Clin Epidemiol. 2007 Dec;60(12):1298-305. PMID: 17998085.

158. Cuello Garcia CA, Pacheco Alvarado KP, Perez Gaxiola G. Grading recommendations in clinical practice guidelines: randomised experimental evaluation of four different systems. Arch Dis Child. 2011 Aug;96(8):723-8. PMID: 21596725.

159. Shakespeare TP, Gebski V, Tang J, et al. Influence of the way results are presented on research interpretation and medical decision making: the PRIMER collaboration randomized studies. Med Decis Making. 2008 Jan-Feb;28(1):127-37. PMID: 18083993.

160. Kuhn KM, Budescu DV, Hershey JR, et al. Attribute tradeoffs in low probability/high consequence risks: The joint effects of dimension preference and vagueness. Risk Decision and Policy. 1999;4(1):31-46.

161. McAuliff BD, Kovera MB, Nunez G. Can jurors recognize missing control groups, confounds, and experimenter bias in psychological science? Law Hum Behav. 2009 Jun;33(3):247-57. PMID: 18587635.

162. Sheridan SL, Halpern DJ, Viera AJ, et al. Interventions for individuals with low health literacy: a systematic review. J Health Commun. 2011;16 Suppl 3(Supplement 3):30-54. PMID: 21951242.

163. Balas EA, Boren SA. Managing Clinical Knowledge for Health Care Improvement. In: Bemmel J, McCray A, eds. Yearbook of Medical Informatics: patient-centered systems. Stuttgart, German: Schatteur; 2000:65-70.

164. Glasgow R, Klesges L, Dzewaltowski D, et al. The future of health behavior change research: What is needed to improve translation of research into health promotion practice? . Ann Behav Med. 2004;27(1):3-12.

165. Kreuter MW, Bernhardt JM. Reframing the dissemination challenge: a marketing and distribution perspective. Am J Public Health. 2009 Dec;99(12):2123-7. PMID: 19833993.

166. Bowen DJ, Sorenson G, Weiner BJ, et al. Dissemination research in cancer control: where are we and where should we go? . Cancer Causes Control. 2009 Aug;20(4):473-85. PMID: 19224380.

167. Timbie JW, Fox DS, Van Busum K, et al. Five reasons that many comparative effectiveness studies fail to change patient care and clinical practice. Health Aff (Millwood). 2012 Oct;31(10):2168-75. PMID: 23048092.

168. Avorn J, Fischer M. 'Bench to behavior': translating comparative effectiveness research into improved clinical practice. Health Aff (Millwood). 2010 Oct;29(10):1891-900. PMID: 20921491.

Appendix A. Search Strategies

May 18 2012 PsycInfo communication (KQ1) and uncertainty (KQ3) final searches

RCTs+Experimental study types: Uncertainty line #47 (2 results)

RCTs for Communication Line #16 (338 results)

SRs/MAs: Communication = 33 (line 41)

SR/MAs: Uncertainty = 0

Search ID#	Search Terms	Search Options	Actions
S47	S35 or S46	Search modes - Boolean/Phrase	(2)
S46	S34 and S45	Search modes - Boolean/Phrase	(1)
S45	S42 or S43 or S44	Search modes - Boolean/Phrase	(12778)
S44	"controlled clinical trial" OR "controlled clinical trials"	Search modes - Boolean/Phrase	(1816)
S43	"cross-over study" OR "cross-over studies"	Search modes - Boolean/Phrase	(418)
S42	"comparative study" OR "comparative studies"	Search modes - Boolean/Phrase	(10556)
S41	S37 or S40	Search modes - Boolean/Phrase	(33)
S40	S11	Limiters - Methodology: -Systematic Review, -Meta Analysis Search modes - Boolean/Phrase	(24)
S39	S34	Limiters - Methodology: -Systematic Review, -Meta Analysis Search modes - Boolean/Phrase	(0)
S38	S34 AND S36	Limiters - Language: English; Age Groups: Adulthood (18 yrs & older); Population Group: Human Search modes - Boolean/Phrase	(0)
S37	S36 and S11	Search modes - Boolean/Phrase	(28)
S36	(DE "Literature Review" and systematic) OR "systematic review" OR (review AND systematic) OR "meta-analysis"	Search modes - Boolean/Phrase	(24095)
S35	S34 AND S15	Limiters - Language: English; Age Groups: Adulthood (18 yrs & older); Population Group: Human Search modes - Boolean/Phrase	(2)
S34	S33	Limiters - Language: English; Age Groups: Adulthood (18 yrs & older); Population Group: Human Search modes - Boolean/Phrase	(249)
S33	S31 and S32	Search modes - Boolean/Phrase	(754)
S32	DE "Communication"	Search modes - Boolean/Phrase	(14325)
S31	S17 or S18 or S19 or S20 or S21 or S22 or S23 or S24 or S25 or S26 or S27 or S28 or S29 or S30	Search modes - Boolean/Phrase	(93893)
S30	"low evidence"	Search modes - Boolean/Phrase	(18)
S29	"conflicting evidence"	Search modes - Boolean/Phrase	(387)
S28	"missing evidence"	Search modes - Boolean/Phrase	(14)
S27	"strength of evidence"	Search modes - Boolean/Phrase	(241)
S26	equipoise	Search modes - Boolean/Phrase	(83)
S25	uncertainty	Search modes - Boolean/Phrase	(17611)
S24	ambigu*	Search modes - Boolean/Phrase	(19891)
S23	complexity	Search modes - Boolean/Phrase	(47888)
S22	vagueness	Search modes - Boolean/Phrase	(627)
S21	precision	Search modes - Boolean/Phrase	(7110)
S20	"net benefit"	Search modes - Boolean/Phrase	(168)

Search ID#	Search Terms	Search Options	Actions
S19	"risk of bias"	Search modes - Boolean/Phrase	(92)
S18	DE "Statistical Reliability"	Search modes - Boolean/Phrase	(3124)
S17	DE "Uncertainty"	Search modes - Boolean/Phrase	(3918)
S16	S15 AND S10	Search modes - Boolean/Phrase	(338)
S15	S12 OR S14	Search modes - Boolean/Phrase	(20758)
S14	control* AND (random* AND trial*)	Search modes - Boolean/Phrase	(20133)
S13	S10 and S12	Search modes - Boolean/Phrase	(192)
S12	"randomized controlled trial" OR RCT*	Search modes - Boolean/Phrase	(7665)
S11	S9 and S10	Search modes - Boolean/Phrase	(3711)
S10	S8	Limiters - Language: English; Age Groups: Adulthood (18 yrs & older); Population Group: Human Search modes - Boolean/Phrase	(15971)
S9	Intervention OR control OR (random* AND trial*)	Limiters - Publication Year from: 2000-2012 Search modes - Boolean/Phrase	(322632)
S8	S1 or S2 or S3 or S4 or S5 or S6	Limiters - Publication Year from: 2000-2012 Search modes - Boolean/Phrase	(34132)
S7	S1 or S2 or S3 or S4 or S5 or S6	Search modes - Boolean/Phrase	(54096)
S6	MM "Health Education"	Search modes - Boolean/Phrase	(7077)
S5	DE "Persuasive Communication"	Search modes - Boolean/Phrase	(4204)
S4	MM "Narratives"	Search modes - Boolean/Phrase	(8157)
S3	DE "Client Education"	Search modes - Boolean/Phrase	(2730)
S2	MM "Decision Making"	Search modes - Boolean/Phrase	(31861)
S1	MM "Information Dissemination"	Search modes - Boolean/Phrase	(617)

May 19-20 2012 PubMed Final CEDT searches: KQ1, KQ2 and KQ3

Summary:

KQ1: RCTs = 3981

SRs+MAs(non-RCTs) = 165

KQ2: RCTs = 1555, after duplicates with KQ1 removed = 401; 1154 imported to duplicates library

SRs+MAs (non-RCTs) = 91, after duplicates with KQ1 removed = 61; 30 imported to duplicates library

KQ3 done 5-20 to expand to all experimental study types per Stacey:

Experimental Study Types including RCTs = 577, after duplicates with KQ1 and/or KQ2 removed= 551; 26 imported to duplicates library

SRs+MAs (Non-RCTs)= 17, after duplicates with KQ1 and/or KQ2 removed = 14; 3 imported to duplicates library

Total from PubMed RCTs (& exp studies of KQ3) = 4933 + 1180 in Duplicates library = 6113

Total from PubMed SR+MA = 240 + 33 in Duplicates library = 275

KQ1 (Communication) search:

Search	Query	Items found
#1	Search "Information Dissemination/methods"[Majr]	1437
#2	Search "Decision Making"[Majr]	44117
#3	Search "Patient Education as Topic"[Mesh]	63946
#4	Search "Narration"[Majr]	1402
#5	Search "Persuasive Communication"[Mesh]	2596
#6	Search "Health Education/methods"[Majr]	14580
#7	Search #1 or #2 or #3 or #4 or #5 or #6	118367
#8	Search #1 or #2 or #3 or #4 or #5 or #6 Filters: Humans	109909
#9	Search #1 or #2 or #3 or #4 or #5 or #6 Filters: Humans; English	98287
#10	Search #1 or #2 or #3 or #4 or #5 or #6 Filters: Humans; English; Adult: 19+ years	37365
#11	Search #1 or #2 or #3 or #4 or #5 or #6 Filters: Publication date from 2000/01/01; Humans; English; Adult: 19+ years	25606
#12	Search #1 or #2 or #3 or #4 or #5 or #6 Filters: Publication date from 2000/01/01; Humans; Comment; English; Adult: 19+ years	335
#13	Search #1 or #2 or #3 or #4 or #5 or #6 Filters: Publication date from 2000/01/01; Humans; Comment; Editorial; English; Adult: 19+ years	448
#14	Search #1 or #2 or #3 or #4 or #5 or #6 Filters: Publication date from 2000/01/01; Humans; Comment; Editorial; Letter; English; Adult: 19+ years	687
#15	Search #1 or #2 or #3 or #4 or #5 or #6 Filters: Publication date from 2000/01/01; Humans; Comment; Editorial; Letter; News; English; Adult: 19+ years	767
#16	Search #11 NOT #15	24839
#17	Search "Randomized Controlled Trial"[Publication Type] OR "Randomized Controlled Trials as Topic"[Mesh] OR "Single-Blind Method"[Mesh] OR "Double-Blind Method"[Mesh] OR "Random Allocation"[Mesh]	476184
#18	Search #16 and #17	3981
#19	Search ("review"[Publication Type] AND "systematic"[tiab]) OR "systematic review"[All Fields] OR ("review literature as topic"[MeSH AND "systematic"[tiab]) OR "meta-analysis"[Publication Type] OR "meta-analysis as topic"[MeSH Terms] OR "meta-analysis"[All Fields]	91360
#20	Search #16 and #19	243
#21	Search #20 NOT #18	165

May 19th 2012 PubMed KQ2 (Dissemination) search:

Search	Query	Items found
#1	Search "Information Dissemination"[Mesh]	7921
#2	Search "Diffusion of Innovation"[Mesh]	13481
#3	Search "Evidence-Based Medicine/education"[Mesh]	1104
#4	Search "Information Services/utilization"[Mesh]	1279
#5	Search "Practice Guidelines as Topic/standards"[Mesh]	4292
#6	Search "Guideline Adherence/statistics and numerical data"[Mesh]	2311
#7	Search "Patient Education as Topic/methods"[Mesh]	11427
#8	Search "Physician's Practice Patterns/standards"[Mesh]	3308
#9	Search "Physician's Practice Patterns/statistics and numerical data"[Mesh]	8826
#10	Search "Physician's Practice Patterns/trends"[Mesh]	2846
#11	Search "Social Marketing"[Mesh]	1756
#12	Search "social marketing"[tiab]	1106
#13	Search "academic detailing"[tiab]	274
#14	Search "dissemination strategy"[tiab]	72
#15	Search "dissemination strategies"[tiab]	125
#16	Search disseminat*[ti] AND guideline*[ti]	119
#17	Search #1 or #2 or #3 or #4 or #5 or #6 or #7 or #8 or #9 or #10 or #11 or #12 or #13 or #14 or #15 or #16	56082
#18	Search #1 or #2 or #3 or #4 or #5 or #6 or #7 or #8 or #9 or #10 or #11 or #12 or #13 or #14 or #15 or #16 Filters: Humans	48743
#19	Search #1 or #2 or #3 or #4 or #5 or #6 or #7 or #8 or #9 or #10 or #11 or #12 or #13 or #14 or #15 or #16 Filters: Humans; English	44665
#20	Search #1 or #2 or #3 or #4 or #5 or #6 or #7 or #8 or #9 or #10 or #11 or #12 or #13 or #14 or #15 or #16 Filters: Humans; English; Adult: 19+ years	13841
#21	Search #1 or #2 or #3 or #4 or #5 or #6 or #7 or #8 or #9 or #10 or #11 or #12 or #13 or #14 or #15 or #16 Filters: Publication date from 2000/01/01; Humans; English; Adult: 19+ years	11150
#22	Search #1 or #2 or #3 or #4 or #5 or #6 or #7 or #8 or #9 or #10 or #11 or #12 or #13 or #14 or #15 or #16 Filters: Publication date from 2000/01/01; Humans; Comment; English; Adult: 19+ years	151
#23	Search #1 or #2 or #3 or #4 or #5 or #6 or #7 or #8 or #9 or #10 or #11 or #12 or #13 or #14 or #15 or #16 Filters: Publication date from 2000/01/01; Humans; Comment; Editorial; English; Adult: 19+ years	204
#24	Search #1 or #2 or #3 or #4 or #5 or #6 or #7 or #8 or #9 or #10 or #11 or #12 or #13 or #14 or #15 or #16 Filters: Publication date from 2000/01/01; Humans; Comment; Editorial; Letter; English; Adult: 19+ years	324
#25	Search #1 or #2 or #3 or #4 or #5 or #6 or #7 or #8 or #9 or #10 or #11 or #12 or #13 or #14 or #15 or #16 Filters: Publication date from 2000/01/01; Humans; Comment; Editorial; Letter; News; English; Adult: 19+ years	362
#26	Search #21 NOT #25	10788
#27	Search "Randomized Controlled Trial"[Publication Type] OR "Randomized Controlled Trials as Topic"[Mesh] OR "Single-Blind Method"[Mesh] OR "Double-Blind Method"[Mesh] OR "Random Allocation"[Mesh]	476184
#28	Search #26 and #27	1555
#29	Search ("review"[Publication Type] AND "systematic"[tiab]) OR "systematic review"[All Fields] OR ("review literature as topic"[MeSH AND "systematic"[tiab]) OR "meta-analysis"[Publication Type] OR "meta-analysis as topic"[MeSH Terms] OR "meta-analysis"[All Fields]	91360
#30	Search #26 and #29	121
#31	Search #30 NOT #28	91

May 20th 2012 PubMed KQ3 (Uncertainty) search to expand to experimental study types:

Search	Query	Items found
#1	Search "Uncertainty"[Mesh]	4566
#2	Search "Research Design/statistics and numerical data"[Mesh]	1286
#3	Search "Therapeutic Equipoise"[Mesh]	58
#4	Search "Bias (Epidemiology)"[Mesh]	44913
#5	Search "low evidence"	66
#6	Search "conflicting evidence"	1594
#7	Search "missing evidence"	52
#8	Search "strength of evidence"	866
#9	Search uncertainty	35336
#10	Search ambigu*	20766
#11	Search complexity	66735
#12	Search vagueness	302
#13	Search precision	64453
#14	Search "risk of bias"	1551
#15	Search "net benefit"	803
#16	Search #1 or #2 or #3 or #4 or #5 or #6 or #7 or #8 or #9 or #10 or #11 or #12 or #13 or #14 or #15	232698
#17	Search #1 or #2 or #3 or #4 or #5 or #6 or #7 or #8 or #9 or #10 or #11 or #12 or #13 or #14 or #15 Filters: Humans	151010
#18	Search #1 or #2 or #3 or #4 or #5 or #6 or #7 or #8 or #9 or #10 or #11 or #12 or #13 or #14 or #15 Filters: Humans; English	139800
#19	Search #1 or #2 or #3 or #4 or #5 or #6 or #7 or #8 or #9 or #10 or #11 or #12 or #13 or #14 or #15 Filters: Humans; English; Adult: 19+ years	48967
#20	Search #1 or #2 or #3 or #4 or #5 or #6 or #7 or #8 or #9 or #10 or #11 or #12 or #13 or #14 or #15 Filters: Humans; Comment; English; Adult: 19+ years	502
#21	Search #1 or #2 or #3 or #4 or #5 or #6 or #7 or #8 or #9 or #10 or #11 or #12 or #13 or #14 or #15 Filters: Humans; Comment; Editorial; English; Adult: 19+ years	558
#22	Search #1 or #2 or #3 or #4 or #5 or #6 or #7 or #8 or #9 or #10 or #11 or #12 or #13 or #14 or #15 Filters: Humans; Comment; Editorial; Letter; English; Adult: 19+ years	688
#23	Search #1 or #2 or #3 or #4 or #5 or #6 or #7 or #8 or #9 or #10 or #11 or #12 or #13 or #14 or #15 Filters: Humans; Comment; Editorial; Letter; News; English; Adult: 19+ years	701
#24	Search #19 NOT #23	48266
#25	Search "Communication"[Mesh]	332362
#26	Search #24 and #25	3172
#27	Search "Randomized Controlled Trial"[Publication Type] OR "Randomized Controlled Trials as Topic"[Mesh] OR "Single-Blind Method"[Mesh] OR "Double-Blind Method"[Mesh] OR "Random Allocation"[Mesh]	476184
#28	Search #26 and #27	194
#29	Search "Comparative Study" [Publication Type] OR "Controlled Clinical Trial" [Publication Type] OR "Cross-Over Studies"[Mesh]	1641835
#30	Search #26 and #29	423
#31	Search #28 or #30	577
#32	Search ("review"[Publication Type] AND "systematic"[tiab]) OR "systematic review"[All Fields] OR ("review literature as topic"[MeSH AND "systematic"[tiab]) OR "meta-analysis"[Publication Type] OR "meta-analysis as topic"[MeSH Terms] OR "meta-analysis"[All Fields]	91379
#33	Search #26 and #32	33
#34	Search #33 NOT #31	17

May 20 2012 Cochrane Library and Cochrane Trials Registry searches

Cochrane Library KQ1 search:

ID	Search	Hits
#1	"Information Dissemination/methods"[Majr]	34
#2	"Decision Making"[Majr]	4713
#3	"Patient Education as Topic"[Mesh]	5620
#4	"Narration"[Majr]	93
#5	"Persuasive Communication"[Mesh]	212
#6	"Health Education/methods"[Majr]	957
#7	(#1 OR #2 OR #3 OR #4 OR #5 OR #6)	11230
#8	(#7), from 2000 to 2012	8489
#9	"Randomized Controlled Trial"[Publication Type] OR "Randomized Controlled Trials as Topic"[Mesh] OR "Single-Blind Method"[Mesh] OR "Double-Blind Method"[Mesh] OR "Random Allocation"[Mesh]	363212
#10	(#8 AND #9)	6161
#11	("review"[Publication Type] AND "systematic"[tiab]) OR "systematic review"[All Fields] OR ("review literature as topic"[MeSH AND "systematic"[tiab]) OR "meta-analysis"[Publication Type] OR "meta-analysis as topic"[MeSH Terms] OR "meta-analysis"[All Fields]	34619
#12	(#8 AND #11)	1707
#13	(#12 AND NOT #10)	828

Cochrane Library KQ2 search:

ID	Search	Hits
#1	"Information Dissemination"[Mesh]	218
#2	"Diffusion of Innovation"[Mesh]	124
#3	"Evidence-Based Medicine/education"[Mesh]	51
#4	"Information Services/utilization"[Mesh]	9
#5	"Practice Guidelines as Topic/standards"[Mesh]	81
#6	"Guideline Adherence/statistics and numerical data"[Mesh]	0
#7	"Patient Education as Topic/methods"[Mesh]	1855
#8	"Physician's Practice Patterns/standards"[Mesh]	93
#9	"Physician's Practice Patterns/statistics and numerical data"[Mesh]	0
#10	"Physician's Practice Patterns/trends"[Mesh]	16
#11	"Social Marketing"[Mesh]	193
#12	"social marketing"[tiab]	193
#13	"academic detailing"[tiab]	164
#14	"dissemination strategy"[tiab]	22
#15	"dissemination strategies"[tiab]	31
#16	disseminat*[ti] AND guideline*[ti]	3685
#17	(#1 OR #2 OR #3 OR #4 OR #5 OR #6 OR #7 OR #8 OR #9 OR #10 OR #11 OR #12 OR #13 OR #14 OR #15 OR #16)	6233
#18	(#17), from 2000 to 2012	5264
#19	"Randomized Controlled Trial"[Publication Type] OR "Randomized Controlled Trials as Topic"[Mesh] OR "Single-Blind Method"[Mesh] OR "Double-Blind Method"[Mesh] OR "Random Allocation"[Mesh]	363212
#20	(#18 AND #19)	2786
#21	("review"[Publication Type] AND "systematic"[tiab]) OR "systematic review"[All Fields] OR ("review literature as topic"[MeSH AND "systematic"[tiab]) OR "meta-analysis"[Publication Type] OR "meta-analysis as topic"[MeSH Terms] OR "meta-analysis"[All Fields]	34619
#22	(#18 AND #21)	2523
#23	(#22 AND NOT #20)	1501

Cochrane Library KQ3 search:

ID	Search	Hits
#1	"Uncertainty"[Mesh]	6758
#2	"Research Design/statistics and numerical data"[Mesh]	0
#3	"Therapeutic Equipoise"[Mesh]	1
#4	"Bias (Epidemiology)"[Mesh]	543
#5	"low evidence"	24
#6	"conflicting evidence"	344
#7	"missing evidence"	5
#8	"strength of evidence"	476
#9	uncertainty	7662
#10	ambigu*	737
#11	complexity	1542
#12	vagueness	13
#13	precision	2621
#14	"risk of bias"	7335
#15	"net benefit"	352
#16	(#1 OR #2 OR #3 OR #4 OR #5 OR #6 OR #7 OR #8 OR #9 OR #10 OR #11 OR #12 OR #13 OR #14 OR #15)	17326
#17	"Communication"[Mesh]	7106
#18	(#16 AND #17)	2144
#19	"Randomized Controlled Trial"[Publication Type] OR "Randomized Controlled Trials as Topic"[Mesh] OR "Single-Blind Method"[Mesh] OR "Double-Blind Method"[Mesh] OR "Random Allocation"[Mesh]	363212
#20	(#18 AND #19)	1776
#21	"Comparative Study" OR "Controlled Clinical Trial" OR "Cross-Over Studies"[Mesh]	224908
#22	(#18 AND #21)	1273
#23	(#20 OR #22)	1837
#24	("review"[Publication Type] AND "systematic"[tiab]) OR "systematic review"[All Fields] OR ("review literature as topic"[MeSH AND "systematic"[tiab]) OR "meta-analysis"[Publication Type] OR "meta-analysis as topic"[MeSH Terms] OR "meta-analysis"[All Fields]	34619
#25	(#18 AND #24)	1980
#26	(#25 AND NOT #23)	231

Web of Science Searches forward-tracing (citations counts) for key reviews, KQ1 and KQ3, May 21-22, 2012

KQ1 Web of Science searches

Kühberger A. The Influence of Framing on Risky Decisions: A Meta-analysis.
Organ Behav Hum Decis Process. 1998 Jul;75(1):23-55. PubMed PMID: 9719656.
Cited by = 219, 218 imported to SR+MA database, 1 dup

Noar, S. M., **Harrington, N. G.**, Van Stee, S. K., & Aldrich, R. S. (2011). Tailored health communication to change lifestyle behaviors. *American Journal of Lifestyle Medicine, 5*(2), 112-122.
Cited by = not in WoS. Cited by 8 in Google Scholar. All imported

Allen, M., & Preiss, R. W. (1997). Comparing the persuasiveness of narrative and statistical evidence using meta-analysis. *Communication Research Reports*, *14*, 125-133.
Cited by = not in WoS, Searched Google Scholar and downloaded 81 references. *All imported*

Title: A meta-analysis of computer-tailored interventions for health behavior change
Author(s): Krebs Paul ; Prochaska James O. ; Rossi Joseph S.
Source: PREVENTIVE MEDICINE Volume: 51 Issue: 3-4 Pages: 214-221 DOI: 10.1016/j.ypmed.2010.06.004 Published: SEP-OCT 2010
Cited by= 15, all imported.

Title: A systematic review of three approaches for constructing physical activity messages: What messages work and what improvements are needed?
Author(s): Latimer Amy E. ; Brawley Lawrence R. ; Bassett Rebecca L.
Source: INTERNATIONAL JOURNAL OF BEHAVIORAL NUTRITION AND PHYSICAL ACTIVITY Volume: 7 Article Number: 36 DOI: 10.1186/1479-5868-7-36 Published: MAY 11 2010
Cited by = 12, 8 imported, 4 dups discarded

Title: Does tailoring matter? Meta-analytic review of tailored print health Behavior change interventions
Author(s): Noar Seth M. ; Benac Christina N. ; Harris Melissa S.
Source: PSYCHOLOGICAL BULLETIN Volume: 133 Issue: 4 Pages: 673-693 DOI: 10.1037/0033-2909.133.4.673 Published: JUL 2007
Cited by = 191, 186 imported, 5 discarded

Title: Does narrative information bias individual's decision making? A systematic review
Author(s): Winterbottom Anna ; Bekker Hilary L. ; Conner Mark ; et al.
Source: SOCIAL SCIENCE & MEDICINE Volume: 67 Issue: 12 Pages: 2079-2088 DOI: 10.1016/j.socscimed.2008.09.037 Published: DEC 2008
Cited by= 20, 18 imported, 2 deleted

O'Keefe , DJ & Jensen , JD (2006). The advantages of compliance or the disadvantages of noncompliance? A meta-analytic review of the relative persuasive effectiveness of gain-framed and loss-framed messages . Communication Yearbook , 30 , 1 – 43
Cited by = Not in WoS not PubMed. Cited by 72 in Google Scholar. 68 imported. 4 dups discarded

Title: The relative persuasiveness of gain-framed and loss-framed messages for encouraging disease prevention behaviors: A meta-analytic review
Author(s): O'Keefe Daniel J. ; Jensen Jakob D.
Source: JOURNAL OF HEALTH COMMUNICATION Volume: 12 Issue: 7 Pages: 623-644 DOI: 10.1080/10810730701615198 Published: OCT-NOV 2007
Cited by = 46, 32 imported, and 14 dups discarded

Title: The Relative Persuasiveness of Gain-Framed and Loss-Framed Messages for Encouraging Disease Detection Behaviors: A Meta-Analytic Review
Author(s): O'Keefe Daniel J. ; Jensen Jakob D.
Source: JOURNAL OF COMMUNICATION Volume: 59 Issue: 2 Pages: 296-316 DOI: 10.1111/j.1460-2466.2009.01417.x Published: JUN 2009
Cited by = 14, 3 imported and 11 dups discarded

Title: What works with men? A systematic review of health promoting interventions targeting men
Author(s): Robertson Lynn M. ; Douglas Flora ; Ludbrook Anne ; et al.
Source: BMC HEALTH SERVICES RESEARCH Volume: 8 Article Number: 141 DOI: 10.1186/1472-6963-8-141 Published: JUL 3 2008
Cited by = 16, 15 imported and 1 dup discarded

KQ3 Web of Science searches:

Politi MC, Han PK, Col NF. Communicating the uncertainty of harms and benefits of medical interventions. Med Decis Making. 2007 Sep-Oct;27(5):681-95. PMID: 17873256.
Cited by = 50, 48 imported to SR+MA database, 2 duplicates

Babrow AS. The many meanings of uncertainty in illness: toward a systematic accounting. Health communication. 10(1): 1-23.
Cited by = 46, 45 imported to SR+MA database, 1 duplicate

Han PK. Varieties of uncertainty in healthcare: a conceptual taxonomy. Medical Decision Making Epub ahead for print Jan 18 2011.
Cited by = not in WoS. Cited by 2 in Google Scholar.

Appendix B. Exclusions

Wrong Publication (protocols, studies published only as abstracts, studies with no original data)

1. Ali R, Alexander KP. Statins for the primary prevention of cardiovascular events in older adults: a review of the evidence (Structured abstract). Am J Geriatr Pharmacother; 2007. p. 52-63.
2. Baishnab E, Karner C. Primary care based clinics for asthma. Cochrane Database of Systematic Reviews: John Wiley & Sons, Ltd; 2012.
3. Bauer MS, McBride L, Williford WO, et al. Collaborative care for bipolar disorder: Part II. Impact on clinical outcome, function, and costs. Psychiatr Serv. 2006 Jul;57(7):937-45. PMID: 16816277.
4. Bekkers MJ, Simpson SA, Dunstan F, et al. Enhancing the quality of antibiotic prescribing in primary care: qualitative evaluation of a blended learning intervention. BMC Fam Pract. 2010;11:34. PMID: 20459655.
5. Blaha MJ, Nasir K, Blumenthal RS. Statin therapy for healthy men identified as "increased risk". JAMA. 2012 Apr 11;307(14):1489-90. PMID: 22496260.
6. Blair MM, Carson DS, Barrington R. Bisphosphonates in the prevention and treatment of glucocorticoid-induced osteoporosis (Structured abstract). J Fam Pract; 2000. p. 839-48.
7. Chasan-Taber L, Marcus BH, Stanek E, 3rd, et al. A randomized controlled trial of prenatal physical activity to prevent gestational diabetes: design and methods. J Womens Health (Larchmt). 2009 Jun;18(6):851-9. PMID: 19514827.
8. Gayet JL, Lithell H. The Study on Cognition and Prognosis in the Elderly (SCOPE): principal results of a randomized double-blind intervention trial [1] (multiple letters). J Hypertens; 2003. p. 1771-2.
9. Kulier R, Khan KS, Gulmezoglu AM, et al. A cluster randomized controlled trial to evaluate the effectiveness of the clinically integrated RHL evidence -based medicine course. Reproductive Health; 2010. p. 8.
10. Loewenstein DA, Arguelles T, Arguelles S, et al. Potential cultural bias in the neuropsychological assessment of the older adult. J Clin Exp Neuropsychol. 1994 Aug;16(4):623-9. PMID: 7962363.
11. Naranbhai V, Abdool Karim Q, Meyer-Weitz A. Interventions to modify sexual risk behaviours for preventing HIV in homeless youth. Cochrane Database of Systematic Reviews: John Wiley & Sons, Ltd; 2011.
12. Quinn CC, Gruber-Baldini AL, Shardell M, et al. Mobile diabetes intervention study: testing a personalized treatment/behavioral communication intervention for blood glucose control. Contemp Clin Trials; 2009. p. 334-46.
13. Thomson H, Thomas S, Sellstrom E, et al. Housing improvements for health and associated socio-economic outcomes. Cochrane Database of Systematic Reviews: John Wiley & Sons, Ltd; 2010.
14. Verweij LM, Proper KI, Weel AN, et al. Design of the Balance@Work project: systematic development, evaluation and implementation of an occupational health guideline aimed at the prevention of weight gain among employees. BMC Public Health; 2009. p. 461.
15. Yusuf S, Anand S. Hormone replacement therapy: a time for pause. CMAJ. 2002 Aug 20;167(4):357-9. PMID: 12197690.
16. Everett T, Bryant A, Griffin Michelle F, et al. Interventions targeted at women to encourage the uptake of cervical screening. Cochrane Database of Systematic Reviews: John Wiley & Sons, Ltd; 2011.
17. Farmer Anna P, Légaré F, Turcot L, et al. Printed educational materials: effects on professional practice and health care outcomes. Cochrane Database of Systematic Reviews: John Wiley & Sons, Ltd; 2008.

18. Flodgren G, Parmelli E, Doumit G, et al. Local opinion leaders: effects on professional practice and health care outcomes. Cochrane Database of Systematic Reviews: John Wiley & Sons, Ltd; 2011.

19. Forsetlund L, Bjørndal A, Rashidian A, et al. Continuing education meetings and workshops: effects on professional practice and health care outcomes. Cochrane Database of Systematic Reviews: John Wiley & Sons, Ltd; 2009.

20. Hill AM, Hill K, Brauer S, et al. Evaluation of the effect of patient education on rates of falls in older hospital patients: description of a randomised controlled trial. BMC Geriatr. 2009;9:14. PMID: 19393046.

21. Barrett HC, Fiddick L. Evolution and risky decisions. Trends Cog Sci, 2000 Jul;4(7):251-2. PMID: WOS:000088091100001.

22. Brainerd CJ, Reyna VF. Fuzzy-trace theory: Dual processes in memory, reasoning, and cognitive neuroscience. Adv Child Dev Behav, Vol 28. 2001;28:41-100. PMID: WOS:000172026800002.

23. Edwards A, Elwyn G, Mulley A. Explaining risks: turning numerical data into meaningful pictures. Br Med J. 2002 Apr;324(7341):827-30. PMID: WOS:000174960300018.

24. Gigerenzer G, Edwards A. Simple tools for understanding risks: from innumeracy to insight. Br Med J. 2003 Sep;327(7417):741-4. PMID: WOS:000185708300028.

25. Guthrie C. Prospect theory, risk preference, and the law. Northwestern University Law Review. 2003 Spr;97(3):1115-63. PMID: WOS:000183052200004.

26. Harrington NG, Noar SM. Reporting standards for studies of tailored interventions (R2). Health Educ Res. 2011.

27. Honka A, Kaipainen K, Hietala H, et al. Rethinking health: ICT enabled services to empower people to manage their health. Biomed Eng. 2011(99):1-.

28. Jusot F, Khlat M. Time preference and risk aversion in educational inequalities in smoking: a population-based study.

29. Kaipainen, K, Honka, A, Saranummi, N. Personalized behavior change support for disease prevention. IEEE. 2011: 880-883.

30. Mass C, Joslyn S, Pyle J, et al. PROBCAST A Web-Based Portal to Mesoscale Probabilistic Forecasts. Bull Am Meteorolog Soc. 2009 Jul;90(7):1009-14. PMID: WOS:000268898100008.

31. Maule AJ, Hodgkinson GP. Heuristics, biases and strategic decision making. Psychologist. 2002 Feb;15(2):68-71. PMID: WOS:000174024600017.

32. Reynolds WW, Nelson RM. Empirical data and the acceptability of research risk - A commentary on the charitable participation standard. Arch Pediatr Adolesc Med. 2008 Jan;162(1):88-90. PMID: WOS:000252176100015.

33. Stanovich KE, Toplak ME, West RF. The development of rational thought: A taxonomy of heuristics and biases. In: Kail RV, ed. Adv Child Dev Behav, Vol 36. Vol. 36. 2008:251-85.

34. Terre L. Optimizing Tailored Health Communications. Am J Lifestyle Med. 2011;5(2):127-9.

35. Wills CE. On the role of framing effects in assessment of health-related utilities. Med Decis Making. 1999 Oct-Dec;19(4):505-6. PMID: WOS:000083056100021.

36. Babrow AS. Colloquy - Developing multiple-process theories of communication. Human Commun Res. 1998 Sep;25(1):152-5. PMID: WOS:000075506400007.

37. Beck SJ. Risk, communication and health psychology. Health Commun. 2007;21(1):97-9. PMID: WOS:000246767900010.

38. Eisenberg EM. Building a mystery: Toward a new theory of communication and identity. J Commun. 2001 Sep;51(3):534-52. PMID: WOS:000170799000006.

39. Fallowfield L. Psychosocial aspects of risk appraisal. Breast Cancer Research. 2008;10. PMID: WOS:000263923200014.

40. Fischhoff B. Applying the science of communication to the communication of science. Climatic Change. 2011 Oct;108(4):701-5. PMID: WOS:000295523600006.

41. Han PK, Klein WM, Arora NK. Varieties of uncertainty in health care: a conceptual taxonomy. Med Decis Making. 2011 Nov-Dec;31(6):828-38. PMID: 22067431.

42. Han PKJ, Klein WMP, Lehman TC, et al. Laypersons' Responses to the Communication of Uncertainty Regarding Cancer Risk Estimates. Med Decis Making. 2009 May;29(3):391-403. PMID: WOS:000266764700016.

43. Hanson JL. Quandaries of Informed Consent. Arch Intern Med. 2009 Nov;169(20):1914-5. PMID: WOS:000271517800014.

44. Haward MF, Lorenz JM. Communicating risk under conditions of uncertainty: not as simple as it may seem. Acta Paediatr. 2011 May;100(5):651-2. PMID: WOS:000289250200007.

45. Mamdani MM. Health advisories: when good intentions go bad. Can Med Assoc J. 2008 Apr;178(8):1025-6. PMID: WOS:000254417900015.

46. Neuner-Jehle S, Senn O, Wegwarth O, et al. How do family physicians communicate about cardiovascular risk? Frequencies and determinants of different communication formats. BMC Fam Pract. 2011 Apr;12PMID: WOS:000289719700001.

47. Strachan PH. What do you say about risk? Heart. 2010 Jan;96(1):7-8. PMID: WOS:000272872600003.

48. Winkler RL. The costs and benefits of vague information. Reliability Engineering and System Safety. 1996;54:127-32.

49. Reyna VF. A theory of medical decision making and health: fuzzy trace theory. Med Decis Making. 2008 Nov-Dec;28(6):850-65. PMID: 19015287.

50. Pang AT, Wittenbrink CM, Lodha SK. Approaches to uncertainty visualization. Visual Computer. 1997;13(8):370-90. PMID: ISI:000071082900003.

51. Babrow AS. Communication and problematIc integration: Understanding diverging probability and value, ambiguity, ambivalence, and impossibility. Communication Theory. 1992;2:95-130.

52. Schunemann HJ, Fretheim A, Oxman AD. Improving the use of research evidence in guideline development: 13. Applicability, transferability and adaptation. Health Res Pol Syst. 2006;4:25.

53. Atkinson P. Training for certainty. Soc Sci Med. 1984;19(9):949-56. PMID: 6515428.

54. Fox RC. The evolution of medical uncertainty. Milbank Mem Fund Q. 1980 Winter;58(1):1-49. PMID: 6903782.

55. Light D, Jr. Uncertainty and control in professional training. J Health Soc Behav. 1979 Dec;20(4):310-22. PMID: 541485.

56. Politi MC, Han PK, Col NF. Communicating the uncertainty of harms and benefits of medical interventions. Med Decis Making. 2007 Sep-Oct;27(5):681-95. PMID: 17873256.

57. Spiegelhalter DJ. Understanding uncertainty. Ann Fam Med. 2008 May-Jun;6(3):196-7. PMID: 18474880.

58. Edwards A, Unigwe S, Elwyn G, et al. Personalised risk communication for informed decision making about entering screening programs. Cochrane Database Syst Rev. 2003;1(1):CD001865. PMID: 12535419.

59. Hoffrage U, Lindsey S, Hertwig R, et al. Medicine. Communicating statistical information. Science. 2000 Dec 22;290(5500):2261-2. PMID: 11188724.

Not Health-Related Evidence (did not include research-based evidence that was health-related)

1. Clayton JM, Butow PN, Tattersall MH, et al. Randomized controlled trial of a prompt list to help advanced cancer patients and their caregivers to ask questions about prognosis and end-of-life care. J Clin Oncol. 2007 Feb 20;25(6):715-23. PMID: 17308275.

2. Katz SJ, Nissan N, Moyer CA. Crossing the digital divide: evaluating online communication between patients and their providers. Am J Manag Care; 2004. p. 593-8.

3. Langham J, Tucker H, Sloan D, et al. Secondary prevention of cardiovascular disease: a randomised trial of training in information management, evidence-based medicine, both or neither: the PIER trial. Br J Gen Pract; 2002. p. 818-24.

4. Marlenga B, Pickett W, Berg RL. Evaluation of an enhanced approach to the dissemination of the North American Guidelines for Children's Agricultural Tasks: a randomized controlled trial. Prev Med. 2002 Aug;35(2):150-9. PMID: 12200100.

5. Matthias MS. Problematic integration in pregnancy and childbirth: contrasting approaches to uncertainty and desire in obstetric and midwifery care. Health Commun. 2009 Jan;24(1):60-70. PMID: 19204859.

6. Pelayo M, Cebrián D, Areosa A, et al. Effects of online palliative care training on knowledge, attitude and satisfaction of primary care physicians. BMC Fam Pract; 2011. p. 37.

7. Pennington L, Roddam H, Burton C, et al. Promoting research use in speech and language therapy: a cluster randomized controlled trial to compare the clinical effectiveness and costs of two training strategies. Clin Rehabil; 2005. p. 387-97.

8. Igou ER, Bless H. On undesirable consequences of thinking: Framing effects as a function of substantive processing. J Behav Decis Making. 2007 Apr;20(2):125-42. PMID: WOS:000245688400002.

9. Jain SP, Mathur P, Maheswaran D. The Influence of Consumers' Lay Theories on Approach/Avoidance Motivation. J Market Res. 2009 Feb;46(1):56-65. PMID: WOS:000262871000008.

11. Kuhberger A, Tanner C. Risky Choice Framing: Task Versions and a Comparison of Prospect Theory and Fuzzy-Trace Theory. J Behav Decis Making. 2010 Jul;23(3):314-29. PMID: WOS:000280020500005.

12. Morss RE, Lazo JK, Demuth JL. Examining the use of weather forecasts in decision scenarios: results from a US survey with implications for uncertainty communication. Meteorol Appl. 2010 Jun;17(2):149-62. PMID: WOS:000279234900006.

13. Patt A. Assessing model-based and conflict-based uncertainty. Global Environmental Change-Human and Policy Dimensions. 2007 Feb;17(1):37-46. PMID: WOS:000245182200006.

14. Roggeveen AL, Grewal D, Gotlieb J. Does the frame of a comparative ad moderate the effectiveness of extrinsic information cues? J Consum Res. 2006 Jun;33(1):115-22. PMID: WOS:000238584600016.

15. McAuliff BD, Kovera MB, Nunez G. Can jurors recognize missing control groups, confounds, and experimenter bias in psychological science? Law Hum Behav. 2009 Jun;33(3):247-57. PMID: 18587635.

Wrong Population (children < 19 years, non-English speaking, incarcerated, Federal and State policymakers)

1. Donze JR, Tercyak KP. The Survivor Health and Resilience Education (SHARE) Program: Development and Evaluation of a Health Behavior Intervention for Adolescent Survivors of Childhood Cancer. J Clin Psychol Med Settings. 2006;13(2):169-76. PMID: 2006-08533-007.

2. Brownson RC, Dodson EA, Stamatakis KA, et al. Communicating evidence-based information on cancer prevention to state-level policy makers. J Natl Cancer Inst. 2011 Feb 16;103(4):306-16. PMID: 21212381.

Wrong Intervention

1. Allen JK, Dennison-Himmelfarb CR, Szanton SL, et al. Community Outreach and Cardiovascular Health (COACH) Trial: a randomized, controlled trial of nurse practitioner/community health worker cardiovascular disease risk reduction in urban community health centers. Circ Cardiovasc Qual Outcomes. 2011 Nov 1;4(6):595-602. PMID: 21953407.
2. Apkon M, Mattera JA, Lin Z, et al. A randomized outpatient trial of a decision-support information technology tool. Arch Intern Med; 2005. p. 2388-94.
3. Badano LP, Di Lenarda A, Bellotti P, et al. Patients with chronic heart failure encountered in daily clinical practice are different from the "typical" patient enrolled in therapeutic trials. Ital Heart J. 2003 Feb;4(2):84-91. PMID: 12762270.
4. Bauer MS, McBride L, Williford WO, et al. Collaborative care for bipolar disorder: part I. Intervention and implementation in a randomized effectiveness trial. Psychiatr Serv. 2006 Jul;57(7):927-36. PMID: 16816276.
5. Berner ES, Baker CS, Funkhouser E, et al. Do local opinion leaders augment hospital quality improvement efforts? A randomized trial to promote adherence to unstable angina guidelines. Med Care. 2003 Mar;41(3):420-31. PMID: 12618645.
6. Borgmeier I, Westenhoefer J. Impact of different food label formats on healthiness evaluation and food choice of consumers: a randomized-controlled study. BMC Public Health. 2009;9:184. PMID: 19523212.
7. Bosworth HB, Olsen MK, Dudley T, et al. Patient education and provider decision support to control blood pressure in primary care: a cluster randomized trial. Am Heart J. 2009 Mar;157(3):450-6. PMID: 19249414.
8. Bradley P, Oterholt C, Herrin J, et al. Comparison of directed and self-directed learning in evidence-based medicine: a randomised controlled trial. Med Educ. 2005 Oct;39(10):1027-35. PMID: 16178830.
9. Brown JB, Shye D, McFarland BH, et al. Controlled trials of CQI and academic detailing to implement a clinical practice guideline for depression. Jt Comm J Qual Improv. 2000 Jan;26(1):39-54. PMID: 10677821.
10. Bullard MJ, Meurer DP, Colman I, et al. Supporting clinical practice at the bedside using wireless technology. Acad Emerg Med; 2004. p. 1186-92.
11. Carpenter RA, Finley C, Barlow CE. Pilot test of a behavioral skill building intervention to improve overall diet quality. J Nutr Educ Behav; 2004. p. 20-4.
12. Clever SL, Ford DE, Rubenstein LV, et al. Primary care patients' involvement in decision-making is associated with improvement in depression. Med Care. 2006 May;44(5):398-405. PMID: 16641657.
13. Cobos A, Vilaseca J, Asenjo C, et al. Cost effectiveness of a clinical decision support system based on the recommendations of the European Society of Cardiology and other societies for the management of hypercholesterolemia: report of a cluster-randomized trial (Structured abstract). Dis Manag Health Outcomes; 2005. p. 421-32.
14. Coenen S, Van Royen P, Michiels B, et al. Optimizing antibiotic prescribing for acute cough in general practice: a cluster-randomized controlled trial. J Antimicrob Chemother. 2004 Sep;54(3):661-72. PMID: 15282232.
15. Coultas D, Frederick J, Barnett B, et al. A randomized trial of two types of nurse-assisted home care for patients with COPD. Chest. 2005 Oct;128(4):2017-24. PMID: 16236850.
16. Dale W, Hemmerich J, Ghini EA, et al. Can induced anxiety from a negative earlier experience influence vascular surgeons' statistical decision-making? A randomized field experiment with an abdominal aortic aneurysm analog. J Am Coll Surg. 2006 Nov;203(5):642-52. PMID: 17084325.
17. Daucourt V, Saillour-Glenisson F, Michel P, et al. A multicenter cluster randomized controlled trial of strategies to improve thyroid function testing. Med Care. 2003 Mar;41(3):432-41. PMID: 12618646.

18. Davis J, Chryssafidou E, Zamora J, et al. Computer-based teaching is as good as face to face lecture-based teaching of evidence based medicine: a randomised controlled trial. BMC Med Educ; 2007. p. 23.

19. DeBisschop M, Robitaille B. Can a patient information sheet reduce antibiotic use in adult outpatients with acute bronchitis? J Fam Pract. 2002 Apr;51(4):381. PMID: 11978266.

20. Eccles M, Grimshaw J, Steen N, et al. The design and analysis of a randomized controlled trial to evaluate computerized decision support in primary care: The COGENT study. Fam Pract; 2000. p. 180-6.

21. Eldredge JD, Carr R, Broudy D, et al. The effect of training on question formulation among public health practitioners: results from a randomized controlled trial. J Med Libr Assoc. 2008 Oct;96(4):299-309. PMID: 18974808.

22. Fagan AA, Arthur MW, Hanson K, et al. Effects of Communities That Care on the adoption and implementation fidelity of evidence-based prevention programs in communities: results from a randomized controlled trial. Prev Sci; 2011. p. 223-34.

23. Fahey T, Montgomery AA, Peters TJ. Randomized trial evaluating the framing of cardiovascular risk and its impact on blood pressure control [ISRCTN87597585]. BMC Health Serv Res. 2001;1:10. PMID: 11686857.

24. Goderis G, Borgermans L, Grol R, et al. Start improving the quality of care for people with type 2 diabetes through a general practice support program: a cluster randomized trial. Diabetes Res Clin Pract. 2010 Apr;88(1):56-64. PMID: 20047770.

25. Hadley J, Kulier R, Zamora J, et al. Effectiveness of an e-learning course in evidence-based medicine for foundation (internship) training. J R Soc Med; 2010. p. 288-94.

26. Halvorsen PA, Wisloff TF, Stovring H, et al. Therapeutic decisions by number needed to treat and survival gains: a cross-sectional survey of lipid-lowering drug recommendations. Br J Gen Pract. 2011 Aug;61(589):e477-83. PMID: 21801540.

27. Hammond A, Bryan J, Hardy A. Effects of a modular behavioural arthritis education programme: a pragmatic parallel-group randomized controlled trial. Rheumatology (Oxford). 2008 Nov;47(11):1712-8. PMID: 18815153.

28. Harewood GC, Hendrick LM. Prospective, controlled assessment of the impact of formal evidence-based medicine teaching workshop on ability to appraise the medical literature. Ir J Med Sci; 2010. p. 91-4.

29. Hatmi ZN, Tahvildari S, Dabiran S, et al. Teaching evidence-based medicine more effectively. Acta Med Iran. 2010 Sep-Oct;48(5):332-6. PMID: 21287468.

30. Heffer RW, Worchel-Prevatt F, Rae WA, et al. The effects of oral versus written instructions on parents' recall and satisfaction after pediatric appointments. J Dev Behav Pediatr. 1997 Dec;18(6):377-82. PMID: 9431499.

31. Hugenholtz NI, Schaafsma FG, Nieuwenhuijsen K, et al. Effect of an EBM course in combination with case method learning sessions: an RCT on professional performance, job satisfaction, and self-efficacy of occupational physicians. Int Arch Occup Environ Health. 2008 Oct;82(1):107-15. PMID: 18386046.

32. Hullett CR. Using functional theory to promote HIV testing: the impact of value-expressive messages, uncertainty, and fear. Health Commun. 2006;20(1):57-67. PMID: 16813489.

33. Jacobson TA, Gutkin SW, Harper CR. Effects of a global risk educational tool on primary coronary prevention: the Atherosclerosis Assessment Via Total Risk (AVIATOR) study. Curr Med Res Opin. 2006 Jun;22(6):1065-73. PMID: 16846539.

34. Johnson CG, Levenkron JC, Suchman AL, et al. Does physician uncertainty affect patient satisfaction? J Gen Intern Med; 1988. p. 144-9.

35. Kim S, Willett LR, Murphy DJ, et al. Impact of an evidence-based medicine curriculum on resident use of electronic resources: a randomized controlled study. J Gen Intern Med; 2008. p. 1804-8.

36. Labarere J, Bosson JL, Sevestre MA, et al. Intervention targeted at nurses to improve venous thromboprophylaxis. Int J Qual Health Care. 2007 Oct;19(5):301-8. PMID: 17726037.

37. Lafata JE, Gunter MJ, Hsu J, et al. Academic detailing to improve laboratory testing among outpatient medication users. Med Care. 2007 Oct;45(10):966-72. PMID: 17890994.

38. Lagerløv P, Loeb M, Andrew M, et al. Improving doctors' prescribing behaviour through reflection on guidelines and prescription feedback: a randomised controlled study. Qual Health Care: QHC; 2000. p. 159-65.

39. Locatelli F, Covic A, Macdougall IC, et al. Effect of computer-assisted European Best Practice Guideline implementation on adherence and target attainment: ORAMA results. J Nephrol; 2009. p. 662-74.

40. Lochman JE, Boxmeyer C, Powell N, et al. Dissemination of the Coping Power program: importance of intensity of counselor training. J Consult Clin Psychol. 2009 Jun;77(3):397-409. PMID: 19485582.

41. Madridejos-Mora R, Amado-Guirado E, Perez-Rodriguez MT. Effectiveness of the combination of feedback and educational recommendations for improving drug prescription in general practice. Med Care. 2004 Jul;42(7):643-8. PMID: 15213488.

42. Mahler HI, Kulik JA, Harrell J, et al. Effects of UV photographs, photoaging information, and use of sunless tanning lotion on sun protection behaviors. Arch Dermatol. 2005 Mar;141(3):373-80. PMID: 15781679.

43. Mazur DJ, Merz JF. Patients' interpretations of verbal expressions of probability: implications for securing informed consent to medical interventions. Behav Sci Law. 1994 Autumn;12(4):417-26. PMID: 10150709.

44. McFarlane JM, Groff JY, O'Brien JA, et al. Secondary prevention of intimate partner violence: a randomized controlled trial. Nurs Res. 2006 Jan-Feb;55(1):52-61. PMID: 16439929.

45. Meredith LS, Jackson-Triche M, Duan N, et al. Quality improvement for depression enhances long-term treatment knowledge for primary care clinicians. J Gen Intern Med; 2000. p. 868-77.

46. Meyer BC, Raman R, Rao R, et al. The STRokE DOC trial technique: 'video clip, drip, and/or ship'. Int J Stroke; 2007. p. 281-7.

47. Meyer J. Performance with tables and graphs: effects of training and a Visual Search Model. Ergonomics. 2000 Nov;43(11):1840-65. PMID: 11105976.

48. Morrison J, Carroll L, Twaddle S, et al. Pragmatic randomised controlled trial to evaluate guidelines for the management of infertility across the primary care-secondary care interface. BMJ; 2001. p. 1282-4.

49. Pollack CE, Mallya G, Polsky D. The impact of consumer-directed health plans and patient socioeconomic status on physician recommendations for colorectal cancer screening. J Gen Intern Med. 2008 Oct;23(10):1595-601. PMID: 18629590.

50. Potter MB, Gildengorin G, Wang Y, et al. Comparative effectiveness of two pharmacy-based colorectal cancer screening interventions during an annual influenza vaccination campaign. J Am Pharm Assoc (2003). 2010 Mar-Apr 1;50(2):181-7. PMID: 20199960.

51. Rebergen DS, Bruinvels DJ, Van Tulder MW, et al. Cost-effectiveness of guideline-based care for workers with mental health problems (Provisional abstract). J Occup Environ Med; 2009. p. 313-22.

52. Robare JF, Bayles CM, Newman AB, et al. The "10 keys" to healthy aging: 24-month follow-up results from an innovative community-based prevention program. Health Educ Behav; 2011. p. 379-88.

53. Robling MR, Houston HL, Kinnersley P, et al. General practitioners' use of magnetic resonance imaging: an open randomized trial comparing telephone and written requests and an open randomized controlled trial of different methods of local guideline dissemination. Clin Radiol; 2002. p. 402-7.

54. Rohrbach LA, Gunning M, Sun P, et al. The Project Towards No Drug Abuse (TND) dissemination trial: implementation fidelity and immediate outcomes. Prev Sci. 2010 Mar;11(1):77-88. PMID: 19757052.

55. Rood E, Bosman RJ, van der Spoel JI, et al. Use of a computerized guideline for glucose regulation in the intensive care unit improved both guideline adherence and glucose regulation. J Am Med Inform Assoc; 2005. p. 172-80.

56. Rosen NO, Knäuper B. A little uncertainty goes a long way: state and trait differences in uncertainty interact to increase information seeking but also increase worry. Health Commun; 2009. p. 228-38.

57. Roumie CL, Elasy TA, Greevy R, et al. Improving blood pressure control through provider education, provider alerts, and patient education: a cluster randomized trial. Ann Intern Med. 2006 Aug 1;145(3):165-75. PMID: 16880458.

58. Roy PM, Durieux P, Gillaizeau F, et al. A computerized handheld decision-support system to improve pulmonary embolism diagnosis: a randomized trial. Ann Intern Med; 2009. p. 677-86.

59. Schilling K, Wiecha J, Polineni D, et al. An interactive web-based curriculum on evidence-based medicine: design and effectiveness. Fam Med; 2006. p. 126-32.

60. Shiffman S, Burton SL, Pillitteri JL, et al. Test of "Light" cigarette counter-advertising using a standard test of advertising effectiveness. Tob Control. 2001;10 Suppl 1:i33-40. PMID: 11740042.

61. Sloman SA, Fernbach PM, Hagmayer Y. Self-deception requires vagueness. Cognition. 2010 May;115(2):268-81. PMID: 20097330.

62. Spoth R, Clair S, Greenberg M, et al. Toward dissemination of evidence-based family interventions: maintenance of community-based partnership recruitment results and associated factors. J Fam Psychol; 2007. p. 137-46.

63. Stevenson K, Lewis M, Hay E. Do physiotherapists' attitudes towards evidence-based practice change as a result of an evidence-based educational programme? J Eval Clin Pract; 2004. p. 207-17.

64. Subramanian U, Fihn SD, Weinberger M, et al. A controlled trial of including symptom data in computer-based care suggestions for managing patients with chronic heart failure. Am J Med. 2004 Mar 15;116(6):375-84. PMID: 15006586.

65. Tamblyn R, Reidel K, Huang A, et al. Increasing the detection and response to adherence problems with cardiovascular medication in primary care through computerized drug management systems: a randomized controlled trial. Med Decis Making. 2010 Mar-Apr;30(2):176-88. PMID: 19675319.

66. Tierney WM, Overhage JM, Murray MD, et al. Can computer-generated evidence-based care suggestions enhance evidence-based management of asthma and chronic obstructive pulmonary disease? A randomized, controlled trial. Health Serv Res. 2005 Apr;40(2):477-97. PMID: 15762903.

67. Trevithick JR, Massel D, Robertson JM, et al. Model study of AREDS antioxidant supplementation of AMD compared to Visudyne: a dominant strategy? (Structured abstract). Ophthalmic Epidemiol; 2004. p. 337-46.

68. Triano JJ, Scaringe J, Bougie J, et al. Effects of visual feedback on manipulation performance and patient ratings. J Manipulative Physiol Ther. 2006 Jun;29(5):378-85. PMID: 16762666.

69. Tsuji SR, Atallah AN, Aranha FC, et al. Cluster randomized clinical trial (ISRCTN23732000) to evaluate the effectiveness of a diagnosis recognition and treatment guide for depressive disorders in primary care. J Eval Clin Pract; 2009. p. 222-5.

70. Vallance JK, Courneya KS, Plotnikoff RC, et al. Maintenance of physical activity in breast cancer survivors after a randomized trial. Med Sci Sports Exerc. 2008 Jan;40(1):173-80. PMID: 18091007.

71. van Roon AH, Hol L, Wilschut JA, et al. Advance notification letters increase adherence in colorectal cancer screening: a population-based randomized trial. Prev Med. 2011 Jun;52(6):448-51. PMID: 21457725.

72. Verhoeven AA, Boerma EJ, Meyboom-de Jong B. Which literature retrieval method is most effective for GPs? Fam Pract. 2000 Feb;17(1):30-5. PMID: 10673485.

73. Verstappen WH, van der Weijden T, Sijbrandij J, et al. Effect of a practice-based strategy on test ordering performance of primary care physicians: a randomized trial. JAMA; 2003. p. 2407-12.

74. Weingarten S, Agocs L, Tankel N, et al. Reducing lengths of stay for patients hospitalized with chest pain using medical practice guidelines and opinion leaders. Am J Cardiol. 1993 Feb 1;71(4):259-62. PMID: 8427164.

75. Wells KB, Sherbourne C, Schoenbaum M, et al. Impact of disseminating quality improvement programs for depression in managed primary care: a randomized controlled trial JAMA 2000 Jun 28;283(24):3204. JAMA; 2000. p. 212-20.

76. Werch CE, Lundstrum RH, Moore A. Bogus-pipeline effects on self-reported college student drug use, problems, and attitudes. Int J Addict. 1989 Oct;24(10):1003-10. PMID: 2620999.

77. Wolf AM, Nasser JF, Schorling JB. The impact of informed consent on patient interest in prostate-specific antigen screening. Arch Intern Med. 1996 Jun 24;156(12):1333-6. PMID: 8651843.

78. Wu AW, Huang IC, Stokes S, et al. Disclosing medical errors to patients: it's not what you say, it's what they hear. J Gen Intern Med. 2009 Sep;24(9):1012-7. PMID: 19578819.

79. Zillich AJ, Ackermann RT, Stump TE, et al. An evaluation of educational outreach to improve evidence-based prescribing in Medicaid: a cautionary tale. J Eval Clin Pract. 2008 Oct;14(5):854-60. PMID: 19018918.

80. Amoako E, Skelly AH. Managing uncertainty in diabetes: an intervention for older African American women. Ethn Dis; 2007. p. 515-21.

81. Amoako E, Skelly AH, Rossen EK. Outcomes of an intervention to reduce uncertainty among African American women with diabetes. West J Nurs Res. 2008 Dec;30(8):928-42. PMID: 18596303.

82. Ansari M, Shlipak MG, Heidenreich PA, et al. Improving guideline adherence: a randomized trial evaluating strategies to increase beta-blocker use in heart failure. Circulation; 2003. p. 2799-804.

83. Ayral X, Gicquere C, Duhalde A, et al. Effects of video information on preoperative anxiety level and tolerability of joint lavage in knee osteoarthritis. Arthritis Rheum. 2002 Aug;47(4):380-2. PMID: 12209483.

84. Bertoni AG, Bonds DE, Chen H, et al. Impact of a multifaceted intervention on cholesterol management in primary care practices: guideline adherence for heart health randomized trial. Arch Intern Med; 2009. p. 678-86.

85. Frost P, Haahr JP, Andersen JH. Reduction of pain-related disability in working populations: a randomized intervention study of the effects of an educational booklet addressing psychosocial risk factors and screening workplaces for physical health hazards. Spine (Phila Pa 1976). 2007 Aug 15;32(18):1949-54. PMID: 17700439.

86. Han PK, Kobrin SC, Klein WM, et al. Perceived ambiguity about screening mammography recommendations: association with future mammography uptake and perceptions. Cancer Epidemiol Biomarkers Prev. 2007 Mar;16(3):458-66. PMID: 17372241.

87. Kelly NR, Huffman LC, Mendoza FS, et al. Effects of a videotape to increase use of poison control centers by low-income and Spanish-speaking families: a randomized, controlled trial. Pediatrics. 2003 Jan;111(1):21-6. PMID: 12509549.

88. Knapp P, Raynor DK, Woolf E, et al. Communicating the risk of side effects to patients: an evaluation of UK regulatory recommendations. Drug Saf. 2009;32(10):837-49. PMID: 19722727.

89. Mishel MH, Germino BB, Belyea M, et al. Moderators of an uncertainty management intervention: for men with localized prostate cancer. Nurs Res; 2003. p. 89-97.

90. Rosen NO, Knauper B, Di Dio P, et al. The impact of intolerance of uncertainty on anxiety after receiving an informational intervention about HPV: a randomised controlled study. Psychol Health. 2010 Jul;25(6):651-68. PMID: 20204959.

91. Stevenson K, Lewis M, Hay E. Does physiotherapy management of low back pain change as a result of an evidence-based educational programme? J Eval Clin Pract; 2006. p. 365-75.

92. Stovring H, Gyrd-Hansen D, Kristiansen IS, et al. Communicating effectiveness of intervention for chronic diseases: what single format can replace comprehensive information? BMC Med Inform Decis Mak; 2008. p. 25.

93. Tsur L, Kozer E, Berkovitch M. The effect of drug consultation center guidance on contraceptive use among women using isotretinoin: a randomized, controlled study. J Womens Health (Larchmt). 2008 May;17(4):579-84. PMID: 18447762.

94. Tziraki C, Graubard BI, Manley M, et al. Effect of training on adoption of cancer prevention nutrition-related activities by primary care practices: results of a randomized, controlled study. J Gen Intern Med; 2000. p. 155-62.

95. Vallance JK, Courneya KS, Plotnikoff RC, et al. Randomized controlled trial of the effects of print materials and step pedometers on physical activity and quality of life in breast cancer survivors. J Clin Oncol. 2007 Jun 10;25(17):2352-9. PMID: 17557948.

96. McKenna K, Leykum LK, McDaniel Jr RR. The role of improvising in patient care. Health Care Manage Rev. 2012.

97. Politi MC, Clark MA, Ombao H, et al. The impact of physicians' reactions to uncertainty on patients' decision satisfaction. J Eval Clin Pract. 2011 Aug;17(4):575-8. PMID: WOS:000293232400006.

98. Reiff M, Bernhardt BA, Mulchandani S, et al. "What does it mean?": Uncertainties in understanding results of chromosomal microarray testing. Genet Med. 2012.

99. Spiegelhalter D, Pearson M, Short I. Visualizing Uncertainty About the Future. Science. 2011 Sep;333(6048):1393-400. PMID: WOS:000294672200029.

100. Vishwanath A. Comparing online information effects - A cross-cultural comparison of online information and uncertainty avoidance. Communication Research. 2003 Dec;30(6):579-98. PMID: WOS:000186499700001.

101. Paul CL, Redman S, Sanson-Fisher RW. A cost-effective approach to the development of printed materials: A randomized controlled trial of three strategies. Health Educ Res. 2004 Dec;19(6):698-706. PMID: 15155585.

102. Spitler JR, Stamm PL, Lineberry TW, et al. Multiple interventions to improve provider adherence to diabetes care guidelines. J Clin Outcomes Manag 2004;11:232-40.

103. Budescu DV, Weinberg S, Wallsten TS. Decisions based on numerically and verbally expressed uncertainties. J Exp Psychol. 1988;14:281-94.

104. Kuhn KM, Budescu DV, Hershey JR, et al. Attribute tradeoffs in low probability/high consequence risks: The joint effects of dimension preference and vagueness. Risk Decision and Policy. 1999;4(1):31-46.

105. Curley SP, Yates JF. The center and range of the probability interval as factors affecting ambiguity preferences. Organ Behav Hum Decis Process. 1985;36:273-87.

106. Einhorn HJ, Hogarth RM. Ambiguity and Uncertainty in Probabilistic Inference. Psychol Rev. 1985;92(4):433-61. PMID: ISI:A1985ASY0700001.

107. Benichou J, Gail MH, Mulvihill JJ. Graphs to estimate an individualized risk of breast cancer. J Clin Oncol. 1996 Jan;14(1):103-10. PMID: 8558184.

108. Peters E, Dieckmann N, Dixon A, et al. Less is more in presenting quality information to consumers. Med Care Res Rev. 2007 Apr;64(2):169-90. PMID: ISI:000245357000003.

109. Budescu DV, Kuhn KM, Kramer KM, et al. Modeling certainty equivalents for imprecise gambles. Organ Behav Hum Decis Process. 2002;88:748-68.

Wrong Comparator (comparisons with usual practice—except for KQ3 when evidence is sparse)

1. Allard J, Hebert R, Rioux M, et al. Efficacy of a clinical medication review on the number of potentially inappropriate prescriptions prescribed for community-dwelling elderly people. CMAJ. 2001 May 1;164(9):1291-6. PMID: 11341138.
2. Aragones A, Schwartz MD, Shah NR, et al. A randomized controlled trial of a multilevel intervention to increase colorectal cancer screening among Latino immigrants in a primary care facility. J Gen Intern Med. 2010 Jun;25(6):564-7. PMID: 20213208.
3. Baer JS, Wells EA, Rosengren DB, et al. Agency context and tailored training in technology transfer: a pilot evaluation of motivational interviewing training for community counselors. J Subst Abuse Treat. 2009 Sep;37(2):191-202. PMID: 19339139.
4. Bemelmans WJ, Broer J, de Vries JH, et al. Impact of Mediterranean diet education versus posted leaflet on dietary habits and serum cholesterol in a high risk population for cardiovascular disease. Public Health Nutr. 2000 Sep;3(3):273-83. PMID: 10979147.
5. Bernstein LM, Chapman GB, Elstein AS. Framing effects in choices between multioutcome life-expectancy lotteries. Med Decis Making. 1999 Jul-Sep;19(3):324-38. PMID: WOS:000081383700011.
6. Bloom JR, Stewart SL, Hancock SL. Breast cancer screening in women surviving Hodgkin disease. Am J Clin Oncol. 2006 Jun;29(3):258-66. PMID: 16755179.
7. Bochicchio GV, Smit PA, Moore R, et al. Pilot study of a web-based antibiotic decision management guide. J Am Coll Surg; 2006. p. 459-67.
8. Bonds DE, Hogan PE, Bertoni AG, et al. A multifaceted intervention to improve blood pressure control: The Guideline Adherence for Heart Health (GLAD) study. Am Heart J. 2009 Feb;157(2):278-84. PMID: 19185634.
9. Boter H, Mistiaen P, Groenewegen I. A randomized trial of a Telephone Reassurance Programme for patients recently discharged from an ophthalmic unit. J Clin Nurs. 2000 Mar;9(2):199-206. PMID: 11111610.
10. Bourgeois FT, Simons WW, Olson K, et al. Evaluation of influenza prevention in the workplace using a personally controlled health record: randomized controlled trial. J Med Internet Res. 2008;10(1):e5. PMID: 18343794.
11. Buller DB, Burgoon M, Hall JR, et al. Using language intensity to increase the success of a family intervention to protect children from ultraviolet radiation: predictions from language expectancy theory. Prev Med; 2000. p. 103-13.
12. Carling CLL, Kristoffersen DT, Oxman AD, et al. The Effect of How Outcomes Are Framed on Decisions about Whether to Take Antihypertensive Medication: A Randomized Trial. PLoS ONE. 2010 Mar;5(3)PMID: WOS:000274997100015.
13. Cherubini P, Rumiati R, Rossi D, et al. Improving attitudes toward prostate examinations by loss-framed appeals. J Appl Soc Psychol. 2005 Apr;35(4):732-44. PMID: WOS:000230372200005.
14. Chiauzzi E, Pujol LA, Wood M, et al. painACTION-back pain: a self-management website for people with chronic back pain. Pain Med; 2010. p. 1044-58.
15. Christensen H, Griffiths KM, Jorm AF. Delivering interventions for depression by using the internet: randomised controlled trial. BMJ. 2004 Jan 31;328(7434):265. PMID: 14742346.
16. Christensen H, Leach LS, Barney L, et al. The effect of web based depression interventions on self reported help seeking: randomised controlled trial [ISRCTN77824516]. BMC Psychiatry. 2006;6:13. PMID: 16595018.
17. Clarkson JE, Young L, Ramsay CR, et al. How to influence patient oral hygiene behavior effectively. J Dent Res. 2009 Oct;88(10):933-7. PMID: 19783802.
18. Corson K, Doak MN, Denneson L, et al. Primary care clinician adherence to guidelines for the management of chronic musculoskeletal pain: results from the study of the effectiveness of a collaborative approach to pain. Pain Med. 2011 Oct;12(10):1490-501. PMID: 21943325.

19. Costanza ME, Luckmann R, Stoddard AM, et al. Using tailored telephone counseling to accelerate the adoption of colorectal cancer screening. Cancer Detect Prev. 2007;31(3):191-8. PMID: 17646058.

20. Croudace T, Evans J, Harrison G, et al. Impact of the ICD-10 Primary Health Care (PHC) diagnostic and management guidelines for mental disorders on detection and outcome in primary care. Cluster randomised controlled trial. Br J Psychiatry. 2003 Jan;182:20-30. PMID: 12509314.

21. Diefenbach MA, Mohamed NE, Butz BP, et al. Acceptability and preliminary feasibility of an Internet/CD-ROM-based education and decision program for early-stage prostate cancer patients: Randomized pilot study. J Med Internet Res. 2012;14(1):262-75. PMID: 2012-05745-022. First Author & Affiliation: Diefenbach, Michael A.

22. Djurcilov S, Kim K, Lermusiaux P, et al. Visualizing scalar volumetric data with uncertainty. Comput Graph. 2002;26:239-48.

23. Druckman JN. Using credible advice to overcome framing effects. J L Econ & Org. 2001 Apr;17(1):62-82. PMID: WOS:000168566200003.

24. Emery J, Morris H, Goodchild R, et al. The GRAIDS Trial: a cluster randomised controlled trial of computer decision support for the management of familial cancer risk in primary care. Br J Cancer; 2007. p. 486-93.

25. Finkelstein JA, Davis RL, Dowell SF, et al. Reducing antibiotic use in children: a randomized trial in 12 practices. Pediatrics; 2001. p. 1-7.

26. Finkelstein JA, Huang SS, Kleinman K, et al. Impact of a 16-community trial to promote judicious antibiotic use in Massachusetts. Pediatrics; 2008. p. e15-23.

27. Fleisher L, Kandadai V, Keenan E, et al. Build it, and will they come? Unexpected findings from a study on a Web-based intervention to improve colorectal cancer screening. J Health Commun. 2012;17(1):41-53. PMID: 22217118.

28. Freemantle N, Johnson R, Dennis J, et al. Sleeping with the enemy? A randomized controlled trial of a collaborative health authority/industry intervention to influence prescribing practice. Br J Clin Pharmacol; 2000. p. 174-9.

29. Freemantle N, Nazareth I, Eccles M, et al. A randomised controlled trial of the effect of educational outreach by community pharmacists on prescribing in UK general practice. Br J Gen Pract; 2002. p. 290-5.

30. Fretheim A, Oxman AD, Havelsrud K, et al. Rational prescribing in primary care (RaPP): a cluster randomized trial of a tailored intervention. PLoS Med. 2006 Jun;3(6):e134. PMID: 16737346.

31. Gadomski A, Ackerman S, Burdick P, et al. Efficacy of the North American guidelines for children's agricultural tasks in reducing childhood agricultural injuries. Am J Public Health; 2006. p. 722-7.

32. Garcia-Retamero R, Cokely ET. Effective Communication of Risks to Young Adults: Using Message Framing and Visual Aids to Increase Condom Use and STD Screening. JExP. 2011 Sep;17(3):270-87. PMID: WOS:000295401500007.

33. Garcia-Retamero R, Galesic M. How to Reduce the Effect of Framing on Messages About Health. J Gen Intern Med. 2010 Dec;25(12):1323-9. PMID: WOS:000284978700014.

34. Gattellari M, Ward JE. Does evidence-based information about screening for prostate cancer enhance consumer decision-making? A randomised controlled trial. J Med Screen. 2003;10(1):27-39. PMID: 12790313.

35. Giles JT, Kennedy DT, Dunn EC, et al. Results of a community pharmacy-based breast cancer risk-assessment and education program. Pharmacotherapy. 2001 Feb;21(2):243-53. PMID: 11213861.

36. Gnjidic D, Le Couteur DG, Abernethy DR, et al. A pilot randomized clinical trial utilizing the drug burden index to reduce exposure to anticholinergic and sedative medications in older people. Ann Pharmacother. 2010 Nov;44(11):1725-32. PMID: 20876826.

37. Goldberg HI, Deyo RA, Taylor VM, et al. Can evidence change the rate of back surgery? A randomized trial of community-based education. Eff Clin Pract; 2001. p. 95-104.

38. Goodwin MA, Zyzanski SJ, Zronek S, et al. A clinical trial of tailored office systems for preventive service delivery. The Study to Enhance Prevention by Understanding Practice (STEP-UP). Am J Prev Med. 2001 Jul;21(1):20-8. PMID: 11418253.

39. Griffiths CE, Taylor H, Collins SI, et al. The impact of psoriasis guidelines on appropriateness of referral from primary to secondary care: a randomized controlled trial. Br J Dermatol; 2006. p. 393-400.

40. Halberg AM, Teigen KH, Fostervold KI. Maximum vs. minimum values: preferences of speakers and listeners for upper and lower limit estimates. Acta Psychol (Amst); 2009. p. 228-39.

41. Ireland DM, Clifton PM, Keogh JB. Achieving the salt intake target of 6 g/day in the current food supply in free-living adults using two dietary education strategies. J Am Diet Assoc. 2010 May;110(5):763-7. PMID: 20430138.

42. Johnson W, Shaya FT, Khanna N, et al. The Baltimore Partnership to Educate and Achieve Control of Hypertension (The BPTEACH Trial): a randomized trial of the effect of education on improving blood pressure control in a largely African American population. J Clin Hypertens (Greenwich). 2011 Aug;13(8):563-70. PMID: 21806766.

43. Katz DA, Muehlenbruch DR, Brown RL, et al. Effectiveness of implementing the agency for healthcare research and quality smoking cessation clinical practice guideline: a randomized, controlled trial. J Natl Cancer Inst. 2004 Apr 21;96(8):594-603. PMID: 15100337.

44. Kerns JW, Krist AH, Woolf SH, et al. Patient perceptions of how physicians communicate during prostate cancer screening discussions: a comparison of residents and faculty. Fam Med; 2008. p. 181-7.

45. Kreuter MW, Sugg-Skinner C, Holt CL, et al. Cultural tailoring for mammography and fruit and vegetable intake among low-income African-American women in urban public health centers. Prev Med. 2005 Jul;41(1):53-62. PMID: ISI:000229651800008.

46. Krist AH, Peele E, Woolf SH, et al. Designing a patient-centered personal health record to promote preventive care. BMC Med Inform Decis Mak. 2011;11:73. PMID: 22115059.

47. Lane DS, Messina CR, Cavanagh MF, et al. A provider intervention to improve colorectal cancer screening in county health centers. Med Care. 2008 Sep;46(9 Suppl 1):S109-16. PMID: 18725822.

48. Lauver DR, Settersten L, Kane JH, et al. Tailored messages, external barriers, and women's utilization of professional breast cancer screening over time. Cancer. 2003 Jun 1;97(11):2724-35. PMID: 12767084.

49. Levine DA, Funkhouser EM, Houston TK, et al. Improving care after myocardial infarction using a 2-year internet-delivered intervention: the Department of Veterans Affairs myocardial infarction-plus cluster-randomized trial. Arch Intern Med. 2011 Nov 28;171(21):1910-7. PMID: 22123798.

50. Lozano P, Finkelstein JA, Carey VJ, et al. A multisite randomized trial of the effects of physician education and organizational change in chronic-asthma care: Health outcomes of the Pediatric Asthma Care Patient Outcomes Research Team II Study. Arch Pediatr Adolesc Med. 2004 Sep;158(9):875-83. PMID: 15351753.

51. Markey P, Schattner P. Promoting evidence-based medicine in general practice-the impact of academic detailing. Fam Pract; 2001. p. 364-6.

52. Martens JD, Winkens RA, van der Weijden T, et al. Does a joint development and dissemination of multidisciplinary guidelines improve prescribing behaviour: a pre/post study with concurrent control group and a randomised trial. BMC Health Serv Res. 2006;6:145. PMID: 17081285.

53. Mayoux-Benhamou A, Giraudet-Le Quintrec JS, Ravaud P, et al. Influence of patient education on exercise compliance in rheumatoid arthritis: a prospective 12-month randomized controlled trial. J Rheumatol. 2008 Feb;35(2):216-23. PMID: 18085742.

54. Mazor KM, Baril J, Dugan E, et al. Patient education about anticoagulant medication: is narrative evidence or statistical evidence more effective? Patient Educ Couns. 2007 Dec;69(1-3):145-57. PMID: 17942268.

55. Mettes TG, van der Sanden WJ, Bronkhorst E, et al. Impact of guideline implementation on patient care: a cluster RCT. J Dent Res; 2010. p. 71-6.

56. Murgaff V, Abraham C, McDermott M. Reducing friday alcohol comsumption among moderate, women drinkers: Evaluation of a brief evidence-based intervention. Alcohol Alcohol; 2007. p. 37-41.

57. Myers RE, Daskalakis C, Cocroft J, et al. Preparing African-American men in community primary care practices to decide whether or not to have prostate cancer screening. J Natl Med Assoc. 2005;97(8):1143-54.

58. O'Brien K, Wright J, Conboy F, et al. The effect of orthodontic referral guidelines: a randomised controlled trial. Br Dent J; 2000. p. 392-7.

59. Ornstein S, Nemeth LS, Jenkins RG, et al. Colorectal cancer screening in primary care: translating research into practice. Med Care. 2010 Oct;48(10):900-6. PMID: 20808257.

60. Perria C, Mandolini D, Guerrera C, et al. Implementing a guideline for the treatment of type 2 diabetics: results of a cluster-randomized controlled trial (C-RCT). BMC Health Serv Res; 2007. p. 79.

61. Powers BJ, Danus S, Grubber JM, et al. The effectiveness of personalized coronary heart disease and stroke risk communication. Am Heart J. 2011 Apr;161(4):673-80. PMID: 21473965.

62. Raynes-Greenow CH, Roberts CL, Nassar N, et al. Do audio-guided decision aids improve outcomes? A randomized controlled trial of an audio-guided decision aid compared with a booklet decision aid for Australian women considering labour analgesia. Health Expect. 2009;12(4):407-16. PMID: 2009-21707-007. First Author & Affiliation: Raynes-Greenow, Camille H.

63. Roumie CL, Grogan EL, Falbe W, et al. A three-part intervention to change the use of hormone replacement therapy in response to new evidence. Ann Intern Med. 2004 Jul 20;141(2):118-25. PMID: 15262667.

64. Scales DC, Dainty K, Hales B, et al. A multifaceted intervention for quality improvement in a network of intensive care units: a cluster randomized trial. JAMA; 2011. p. 363-72.

65. Schneeweiss S, Ratnapalan S. Impact of a multifaceted pediatric sedation course: self-directed learning versus a formal continuing medical education course to improve knowledge of sedation guidelines. CJEM; 2007. p. 93-100.

66. Schnoor M, Meyer T, Suttorp N, et al. Development and evaluation of an implementation strategy for the German guideline on community-acquired pneumonia. Qual Saf Health Care. 2010 Dec;19(6):498-502. PMID: 20388644.

67. Searle J, Grover S, Santin A, et al. Randomised trial of an integrated educational strategy to reduce investigation rates in young women with dysfunctional uterine bleeding. Aust N Z J Obstet Gynaecol. 2002 Oct;42(4):395-400. PMID: 12403289.

68. Shekelle PG, Kravitz RL, Beart J, et al. Are nonspecific practice guidelines potentially harmful? A randomized comparison of the effect of nonspecific versus specific guidelines on physician decision making. Health Serv Res; 2000. p. 1429-48.

69. Smith DH, Christensen DB, Stergachis A, et al. A randomized controlled trial of a drug use review intervention for sedative hypnotic medications. Med Care. 1998 Jul;36(7):1013-21. PMID: 9674619.

70. Søndergaard J, Andersen M, Støvring H, et al. Mailed prescriber feedback in addition to a clinical guideline has no impact: a randomised, controlled trial. Scand J Prim Health Care; 2003. p. 47-51.

71. Thompson A, Sullivan SA, Barley M, et al. The DEBIT trial: an intervention to reduce antipsychotic polypharmacy prescribing in adult psychiatry wards - a cluster randomized controlled trial. Psychol Med. 2008 May;38(5):705-15. PMID: 17825122.

72. Vallance J, Plotnikoff RC, Karvinen KH, et al. Understanding physical activity maintenance in breast cancer survivors. Am J Health Behav. 2010 Mar-Apr;34(2):225-36. PMID: 19814602.

73. Vallance JK, Courneya KS, Plotnikoff RC, et al. Analyzing theoretical mechanisms of physical activity behavior change in breast cancer survivors: results from the activity promotion (ACTION) trial. Ann Behav Med. 2008 Apr;35(2):150-8. PMID: 18347895.

74. van der Sanden WJ, Mettes DG, Plasschaert AJ, et al. Effectiveness of clinical practice guideline implementation on lower third molar management in improving clinical decision-making: a randomized controlled trial. Eur J Oral Sci; 2005. p. 349-54.

75. Welkenhuysen M, Evers-Kiebooms G, d'Ydewalle G. The language of uncertainty in genetic risk communication: framing and verbal versus numerical information. Patient Educ Couns. 2001 May;43(2):179-87. PMID: 11369151.

76. Welkenhuysen M, Evers-Kiebooms G, d'Ydewalle G. The language of uncertainty in genetic risk communication: framing and verbal versus numerical information. Patient Educ Couns. 2001 May;43(2):179-87. PMID: WOS:000169169400008.

77. Williams-Piehota P, Pizarro J, Navarro Silvera SA, et al. Need for cognition and message complexity in motivating fruit and vegetable intake among callers to the cancer information service. Health Commun. 2006;19(1):75-84. PMID: 16519594.

78. Windsor R, Woodby L, Miller T, et al. Effectiveness of smoking cessation and reduction in pregnancy treatment (SCRIPT) methods in Medicaid-supported prenatal care: Trial III. Health Educ Behav. 2011;38(4):412-22. PMID: 2011-16907-009. First Author & Affiliation: Windsor, Richard.

79. Wong SS, Thornton JG, Gbolade B, et al. A randomised controlled trial of a decision-aid leaflet to facilitate women's choice between pregnancy termination methods. BJOG. 2006 Jun;113(6):688-94. PMID: 16709212.

80. Yates K, Pena A. Comprehension of discharge information for minor head injury: a randomised controlled trial in New Zealand. N Z Med J. 2006;119(1239):U2101. PMID: 16912719.

Wrong Outcome (outcomes for target audience did not include awareness of, knowledge about, discussions about, self-efficacy to use, or behavioral intentions to use or apply the evidence; outcomes for patients did not include health-related decisions or behaviors or clinical outcomes; ultimate outcomes for clinicians did not include behavior)

1. Baker R, Fraser RC, Stone M, et al. Randomised controlled trial of the impact of guidelines, prioritized review criteria and feedback on implementation of recommendations for angina and asthma. Br J Gen Pract; 2003. p. 284-91.

2. Riggs NR, Nakawatase M, Pentz MA. Promoting community coalition functioning: effects of Project STEP. Prev Sci; 2008. p. 63-72.

3. Yuksel N, Majumdar SR, Biggs C, et al. Community pharmacist-initiated screening program for osteoporosis: randomized controlled trial. Osteoporos Int. 2010 Mar;21(3):391-8. PMID: 19499272.

4. Ryan GL, Skinner CS, Farrell D, et al. Examining the boundaries of tailoring: The utility of tailoring versus targeting mammography interventions for two distinct populations. Health Educ Res. 2001 Oct;16(5):555-66. PMID: 11675803.

5. Steginga S, Occhipinti S. Decision making about treatment of hypothetical prostate cancer: is deferring a decision an expert-opinion heuristic? J Psychosoc Oncol. 2002;20(3):69-84.

6. Belia S, Fidler F, Williams J, et al. Researchers misunderstand confidence intervals and standard error bars. Psychol Methods. 2005 Dec;10(4):389-96. PMID: 16392994.

Wrong Evidence Source (did not provide evidence that was assembled, reviewed, and presented by a professional organization that was then used to make recommendations)

1. Amsallem E, Kasparian C, Cucherat M, et al. Evaluation of two evidence-based knowledge transfer interventions for physicians. A cluster randomized controlled factorial design trial: the CardioDAS Study. Fundam Clin Pharmacol; 2007. p. 631-41.

2. Azocar F, Cuffel B, Goldman W, et al. The impact of evidence-based guideline dissemination for the assessment and treatment of major depression in a managed behavioral health care organization. J Behav Health Ser Res; 2003. p. 109-18.

3. Funk M, Wutzke S, Kaner E, et al. A multicountry controlled trial of strategies to promote dissemination and implementation of brief alcohol intervention in primary health care: findings of a World Health Organization collaborative study. J Stud Alcohol. 2005 May;66(3):379-88. PMID: 16047527.

4. Hall L, Eccles M, Barton R, et al. Is untargeted outreach visiting in primary care effective? A pragmatic randomized controlled trial. J Public Health Med; 2001. p. 109-13.

5. Romero A, Alonso C, Marín I, et al. [Effectiveness of a multifactorial strategy for implementing clinical guidelines on unstable angina: cluster randomized trial]. Rev Esp Cardiol; 2005. p. 640-8.

6. Schroter S, Jenkins RD, Playle RA, et al. Evaluation of an online interactive Diabetes Needs Assessment Tool (DNAT) versus online self-directed learning: a randomised controlled trial. BMC Med Educ; 2011. p. 35.

7. Seager JM, Howell-Jones RS, Dunstan FD, et al. A randomised controlled trial of clinical outreach education to rationalise antibiotic prescribing for acute dental pain in the primary care setting. Br Dent J. 2006 Aug 26;201(4):217-22; discussion 6. PMID: 16902573.

8. Witt K, Knudsen E, Ditlevsen S, et al. Academic detailing has no effect on prescribing of asthma medication in Danish general practice: a 3-year randomized controlled trial with 12-monthly follow-ups. Fam Pract; 2004. p. 248-53.

9. Cousineau TM, Green TC, Corsini E, et al. Online psychoeducational support for infertile women: a randomized controlled trial. Hum Reprod. 2008 Mar;23(3):554-66. PMID: 18089552.

10. Gordon JS, Andrews JA, Lichtenstein E, et al. Disseminating a smokeless tobacco cessation intervention model to dental hygienists: a randomized comparison of personalized instruction and self-study methods. Health Psychol. 2005 Sep;24(5):447-55. PMID: 16162038.

11. Man-Son-Hing M, O'Connor AM, Drake E, et al. The effect of qualitative vs. quantitative presentation of probability estimates on patient decision-making: a randomized trial. Health Expect. 2002 Sep;5(3):246-55. PMID: 12199663.

12. Mishel MH, Germino BB, Lin L, et al. Managing uncertainty about treatment decision making in early stage prostate cancer: a randomized clinical trial. Patient Educ Couns. 2009 Dec;77(3):349-59. PMID: 19819096.

13. Nolan RP, Upshur RE, Lynn H, et al. Therapeutic benefit of preventive telehealth counseling in the Community Outreach Heart Health and Risk Reduction Trial. Am J Cardiol. 2011 Mar 1;107(5):690-6. PMID: 21215382.

14. Ruf D, Berner M, Kriston L, et al. Cluster-randomized controlled trial of dissemination strategies of an online quality improvement programme for alcohol-related disorders. Alcohol Alcohol. 2010 Jan-Feb;45(1):70-8. PMID: 19889887.

15. Schectman JM, Schroth WS, Verme D, et al. Randomized controlled trial of education and feedback for implementation of guidelines for acute low back pain. J Gen Intern Med. 2003 Oct;18(10):773-80. PMID: 14521638.

16. Lipkus IM, Rimer BK, Halabi S, et al. Can tailored interventions increase mammography use among HMO women? Am J Prev Med. 2000;18(1):1-10.

17. Andersen MR, Yasui Y, Meischke H, et al. The effectiveness of mammography promotion by volunteers in rural communities. Am J Prev Med. 2000 Apr;18(3):199-207. PMID: 10722985.

18. Champion V, Maraj M, Hui S, et al. Comparison of tailored interventions to increase mammography screening in nonadherent older women. Prev Med. 2003 Feb;36(2):150-8. PMID: 12590989.

19. Taplin SH, Barlow WE, Ludman E, et al. Testing reminder and motivational telephone calls to increase screening mammography: A randomized study. J Natl Cancer Inst. 2000 Feb 2;92(3):233-42. PMID: 10655440.

20. Valanis BG, Glasgow RE, Mullooly J, et al. Screening HMO women overdue for both mammograms and Pap tests. Prev Med. 2002 Jan;34(1):40-50. PMID: ISI:000173235300006.

21. Di Noia J, Schwinn TM, Dastur ZA, et al. The relative efficacy of pamphlets, CD-ROM, and the Internet for disseminating adolescent drug abuse prevention programs: an exploratory study. Prev Med. 2003 Dec;37(6 Pt 1):646-53. PMID: 14636798.

22. McDonald MV, Pezzin LE, Feldman PH, et al. Can just-in-time, evidence-based "reminders" improve pain management among home health care nurses and their patients? J Pain Symptom Manage. 2005;29(5):474-88.

23. Oyeku SO, Feldman HA, Ryan K, et al. Primary care clinicians' knowledge and confidence about newborn screening for sickle cell disease: Randomized assessment of educational strategies. J Natl Med Assoc. 2010 Aug;102(8):676-82. PMID: ISI:000281129500001.

24. van Eijk ME, Avorn J, Porsius AJ, et al. Reducing prescribing of highly anticholinergic antidepressants for elderly people: randomised trial of group versus individual academic detailing. BMJ. 2001 Mar 17;322(7287):654-7. PMID: 11250852.

Wrong Country (any country not specified for inclusion)

1. Cuello García CA, Pacheco Alvarado KP, Pérez Gaxiola G. Grading recommendations in clinical practice guidelines: randomised experimental evaluation of four different systems. Arch Dis Child; 2011. p. 723-8.

2. Shakespeare TP, Gebski V, Tang J, et al. Influence of the way results are presented on research interpretation and medical decision making: the PRIMER collaboration randomized studies. Med Decis Making. 2008 Jan-Feb;28(1):127-37. PMID: 18083993.

3. Lewis E, Mayer JA, Slymen D, et al. Disseminating a sun safety program to zoological parks: the effects of tailoring. Health Psychol. 2005 Sep;24(5):456-62. PMID: 16162039.

Wrong Study Design (nonrandomized trials, nonexperimental studies, single pre/post designs, qualitative research)

1. Denton GD, Smith J, Faust J, et al. Comparing the efficacy of staff versus housestaff instruction in an intervention to improve hypertension management. Acad Med; 2001. p. 1257-60.
2. Dobbins M, Robeson P, Ciliska D, et al. A description of a knowledge broker role implemented as part of a randomized controlled trial evaluating three knowledge translation strategies. Implementation Science; 2009. p. 23.
3. Ling BS, Schoen RE, Trauth JM, et al. Physicians encouraging colorectal screening: a randomized controlled trial of enhanced office and patient management on compliance with colorectal cancer screening. Arch Intern Med. 2009 Jan 12;169(1):47-55. PMID: 19139323.
4. Rodriguez-Salvanes F, Novella B, Fernandez Luque MJ, et al. Efficacy of a strategy for implementing a guideline for the control of cardiovascular risk in a primary healthcare setting: the SIRVA2 study a controlled, blinded community intervention trial randomised by clusters. BMC Fam Pract. 2011;12:21. PMID: 21504570.
5. Teigen KH, Brun W. The directionality of verbal probability expressions: Effects on decisions, predictions, and probabilistic reasoning. Organ Behav Hum Decis Process. 1999 Nov;80(2):155-90. PMID: WOS:000083525000004.
6. Frosch DL, Kaplan RM, Felitti VJ. A randomized controlled trial comparing internet and video to facilitate patient education for men considering the prostate specific antigen test. J Gen Intern Med. 2003 Oct;18(10):781-7. PMID: 14521639.
7. Frosch DL, Kaplan RM, Felitti V. The evaluation of two methods to facilitate shared decision making for men considering the prostate-specific antigen test. J Gen Intern Med. 2001 Jun;16(6):391-8. PMID: 11422636.

Wrong Sample Size (N < 100 total individuals in the study)

1. Jaffe-Katz A, Budescu DV, Wallsten TS. Timed magnitude comparisons of numerical and nonnumerical expressions of uncertainty. Mem Cognit. 1989 May;17(3):249-64. PMID: 2725262.
2. Larkey LK, Gonzalez J. Storytelling for promoting colorectal cancer prevention and early detection among Latinos. Patient Educ Couns. 2007 Aug;67(3):272-8. PMID: 17524595.
3. Liaw ST, Sulaiman ND, Barton CA, et al. An interactive workshop plus locally adapted guidelines can improve general practitioners asthma management and knowledge: a cluster randomised trial in the Australian setting. BMC Fam Pract; 2008. p. 22.
4. White M, Michaud G, Pachev G, et al. Randomized trial of problem-based versus didactic seminars for disseminating evidence-based guidelines on asthma management to primary care physicians. J. Contin. Educ. Health Professions; 2004. p. 237-43.
5. Akl EA, Maroun N, Guyatt G, et al. Symbols were superior to numbers for presenting strength of recommendations to health care consumers: a randomized trial. J Clin Epidemiol. 2007 Dec;60(12):1298-305. PMID: 17998085.
6. Ibrekk H, Morgan MJ. Graphical communication of uncertain quantities to nontechnical people. Risk Anal. 1987;7(4):519-29.
7. Kuhn KM. Communicating uncertainty: framing effects on response to vague probabilities. Organ Behav Hum Decis Process. 1997;71(1):55-83.

High Risk of Bias

1. Church TR, Yeazel MW, Jones RM, et al. A randomized trial of direct mailing of fecal occult blood tests to increase colorectal cancer screening. J Natl Cancer Inst. 2004 May 19;96(10):770-80. PMID: 15150305.
2. Daniels EC, Bacon J, Denisio S, et al. Translation squared: improving asthma care for high-disparity populations through a safety net practice-based research network. J Asthma; 2005. p. 499-505.
3. Figueiras A, Sastre I, Tato F, et al. One-to-one versus group sessions to improve prescription in primary care: a pragmatic randomized controlled trial. Med Care; 2001. p. 158-67.
4. Johnson BB, Slovic P. Presenting uncertainty in health risk assessment: Initial studies of its effects on risk perception and trust. Risk Anal. 1995 Aug;15(4):485-94. PMID: 7480948.
5. Manfredi C, Cho YI, Warnecke R, et al. Dissemination strategies to improve implementation of the PHS smoking cessation guideline in MCH public health clinics: experimental evaluation results and contextual factors. Health Educ Res. 2011 Apr;26(2):348-60. PMID: 21398375.
6. Renzi PM, Ghezzo H, Goulet S, et al. Paper stamp checklist tool enhances asthma guidelines knowledge and implementation by primary care physicians. Can Respir J; 2006. p. 193-7.
7. van Driel ML, Coenen S, Dirven K, et al. What is the role of quality circles in strategies to optimise antibiotic prescribing? A pragmatic cluster-randomised controlled trial in primary care. Qual Saf Health Care. 2007 Jun;16(3):197-202. PMID: 17545346.
8. van Steenkiste B, van der Weijden T, Stoffers HE, et al. Improving cardiovascular risk management: a randomized, controlled trial on the effect of a decision support tool for patients and physicians. Eur J Cardiovasc Prev Rehabil. 2007 Feb;14(1):44-50. PMID: 17301626.
9. Varonen H, Rautakorpi UM, Nyberg S, et al. Implementing guidelines on acute maxillary sinusitis in general practice--a randomized controlled trial. Fam Pract; 2007. p. 201-6.
10. Wadland WC, Soffelmayr B, Ives K. Enhancing smoking cessation of low-income smokers in managed care. J Fam Pract. 2001 Feb;50(2):138-44. PMID: 11219562.

Appendix C. Risk of Bias Tables

Table C-1. Risk of bias for KQ 1

Author, Year	Question	Response
Cox, 2001[1]	Q1: Does the study population match the source population defined by the inclusion/exclusion criteria?"	No
	Q2: Was randomization adequately performed?	NR
	Q3: Was allocation adequately concealed?	NR
	Q4: Did randomization result in equal practice and participant characteristics among study groups?	No
	Q5: Did the study maintain comparable groups?	NR
	Q6: In cases of high loss to follow-up (or differential loss to follow-up), is the impact assessed (e.g., through sensitivity analysis or other adjustment method)?	NA
	Q7: Are outcomes assessed using valid and reliable measures?	Fair
	Q8: Are outcomes implemented consistently across all study participants?	Good
	Q9: Were participants and intervention deliverers blinded during the study	Fair
	Q10: Were the outcome assessors (e.g., data collectors) blinded to the intervention arm of participants?	No
	Q11: Was the intervention delivered to all members of the intervention group as described?	NR
	Q12: Are the outcomes of interest pre-specified by the researchers?	Good
	Q13: Were the key confounding variables taken into account in the analysis and/or through randomization (e.g., successful randomization without loss to follow up, through matching, stratification, multivariate analysis, exclude based on confounding or other statistical adjustment)?	Yes
	Q14: Are the statistical methods used to assess the primary benefit outcomes appropriate to the data?	Yes
	Q15: Are results believable taking study limitations into consideration?	Fair

Table C-1. Risk of bias for KQ 1 (continued)

Author, Year	Question	Response
Elder, 2005[2]	Q1: Does the study population match the source population defined by the inclusion/exclusion criteria?"	Yes
	Q2: Was randomization adequately performed?	NR
	Q3: Was allocation adequately concealed?	NR
	Q4: Did randomization result in equal practice and participant characteristics among study groups?	No
	Q5: Did the study maintain comparable groups?	Yes
	Q6: In cases of high loss to follow-up (or differential loss to follow-up), is the impact assessed (e.g., through sensitivity analysis or other adjustment method)?	NA
	Q7: Are outcomes assessed using valid and reliable measures?	Good
	Q8: Are outcomes implemented consistently across all study participants?	Good
	Q9: Were participants and intervention deliverers blinded during the study	Fair
	Q10: Were the outcome assessors (e.g., data collectors) blinded to the intervention arm of participants?	Yes
	Q11: Was the intervention delivered to all members of the intervention group as described?	Good
	Q12: Are the outcomes of interest pre-specified by the researchers?	Good
	Q13: Were the key confounding variables taken into account in the analysis and/or through randomization (e.g., successful randomization without loss to follow up, through matching, stratification, multivariate analysis, exclude based on confounding or other statistical adjustment)?	Yes
	Q14: Are the statistical methods used to assess the primary benefit outcomes appropriate to the data?	Yes
	Q15: Are results believable taking study limitations into consideration?	Fair

Table C-1. Risk of bias for KQ 1 (continued)

Author, Year	Question	Response
Jibaja-Weiss, 2003[3]	Q1: Does the study population match the source population defined by the inclusion/exclusion criteria?"	Yes
	Q2: Was randomization adequately performed?	Yes
	Q3: Was allocation adequately concealed?	Yes
	Q4: Did randomization result in equal practice and participant characteristics among study groups?	Yes
	Q5: Did the study maintain comparable groups?	NR
	Q6: In cases of high loss to follow-up (or differential loss to follow-up), is the impact assessed (e.g., through sensitivity analysis or other adjustment method)?	NA
	Q7: Are outcomes assessed using valid and reliable measures?	Good
	Q8: Are outcomes implemented consistently across all study participants?	Good
	Q9: Were participants and intervention deliverers blinded during the study	Good
	Q10: Were the outcome assessors (e.g., data collectors) blinded to the intervention arm of participants?	Yes
	Q11: Was the intervention delivered to all members of the intervention group as described?	Fair
	Q12: Are the outcomes of interest pre-specified by the researchers?	Good
	Q13: Were the key confounding variables taken into account in the analysis and/or through randomization (e.g., successful randomization without loss to follow up, through matching, stratification, multivariate analysis, exclude based on confounding or other statistical adjustment)?	Yes
	Q14: Are the statistical methods used to assess the primary benefit outcomes appropriate to the data?	Yes
	Q15: Are results believable taking study limitations into consideration?	Fair

Table C-1. Risk of bias for KQ 1 (continued)

Author, Year	Question	Response
Schneider, 2001[4]	Q1: Does the study population match the source population defined by the inclusion/exclusion criteria?"	No
	Q2: Was randomization adequately performed?	Yes
	Q3: Was allocation adequately concealed?	No
	Q4: Did randomization result in equal practice and participant characteristics among study groups?	NR- They don't break out the different intervention groups to see if randomization worked; they just look at breakdown by ethnicity; there's a footnote-3 on p. 258 that does list some numbers but not all
	Q5: Did the study maintain comparable groups?	No
	Q6: In cases of high loss to follow-up (or differential loss to follow-up), is the impact assessed (e.g., through sensitivity analysis or other adjustment method)?	No
	Q7: Are outcomes assessed using valid and reliable measures?	Fair- self-reported receipt of mammo is a Fair measure; it has some issues with reliability but other studies use it and it's Fairly valid
	Q8: Are outcomes implemented consistently across all study participants?	Fair
	Q9: Were participants and intervention deliverers blinded during the study	Fair
	Q10: Were the outcome assessors (e.g., data collectors) blinded to the intervention arm of participants?	NR
	Q11: Was the intervention delivered to all members of the intervention group as described?	Fair- people were assigned to the intervention either individually or at a group level (p 258); I think that means Fair
	Q12: Are the outcomes of interest pre-specified by the researchers?	Good
	Q13: Were the key confounding variables taken into account in the analysis and/or through randomization (e.g., successful randomization without loss to follow up, through matching, stratification, multivariate analysis, exclude based on confounding or other statistical adjustment)?	Yes
	Q14: Are the statistical methods used to assess the primary benefit outcomes appropriate to the data?	Yes
	Q15: Are results believable taking study limitations into consideration?	Fair

Table C-1. Risk of bias for KQ 1 (continued)

Author, Year	Question	Response
Vernon, 2008[5]	Q1: Does the study population match the source population defined by the inclusion/exclusion criteria?"	Yes
	Q2: Was randomization adequately performed?	Yes
	Q3: Was allocation adequately concealed?	Yes
	Q4: Did randomization result in equal practice and participant characteristics among study groups?	Yes
	Q5: Did the study maintain comparable groups?	Yes
	Q6: In cases of high loss to follow-up (or differential loss to follow-up), is the impact assessed (e.g., through sensitivity analysis or other adjustment method)?	Yes
	Q7: Are outcomes assessed using valid and reliable measures?	Good
	Q8: Are outcomes implemented consistently across all study participants?	Good
	Q9: Were participants and intervention deliverers blinded during the study	Fair
	Q10: Were the outcome assessors (e.g., data collectors) blinded to the intervention arm of participants?	Yes
	Q11: Was the intervention delivered to all members of the intervention group as described?	Fair
	Q12: Are the outcomes of interest pre-specified by the researchers?	Good
	Q13: Were the key confounding variables taken into account in the analysis and/or through randomization (e.g., successful randomization without loss to follow up, through matching, stratification, multivariate analysis, exclude based on confounding or other statistical adjustment)?	Yes
	Q14: Are the statistical methods used to assess the primary benefit outcomes appropriate to the data?	Yes
	Q15: Are results believable taking study limitations into consideration?	Good

Table C-2. Risk of bias for KQ 2

Author, Year	Question	Response
Bahrami, 2004[6]	Q1: Does the study population match the source population defined by the inclusion/exclusion criteria?"	Yes
	Q2: Was randomization adequately performed?	Yes
	Q3: Was allocation adequately concealed?	Yes
	Q4: Did randomization result in equal practice and participant characteristics among study groups?	Yes
	Q5: Did the study maintain comparable groups?	Yes
	Q6: In cases of high loss to follow-up (or differential loss to follow-up), is the impact assessed (e.g., through sensitivity analysis or other adjustment method)?	NA
	Q7: Are outcomes assessed using valid and reliable measures?	Good
	Q8: Are outcomes implemented consistently across all study participants?	Good
	Q9: Were participants and intervention deliverers blinded during the study?	Fair
	Q10: Were the outcome assessors (e.g., data collectors) blinded to the intervention arm of participants?	Yes
	Q11: Was the intervention delivered to all members of the intervention group as described?	Fair
	Q12: Are the outcomes of interest pre-specified by the researchers?	Good
	Q13: Were the key confounding variables taken into account in the analysis and/or through randomization (e.g., successful randomization without loss to follow up, through matching, stratification, multivariate analysis, exclude based on confounding or other statistical adjustment)?	Yes
	Q14: Are the statistical methods used to assess the primary benefit outcomes appropriate to the data?	Yes
	Q15: Are results believable taking study limitations into consideration?	Good
Banait, 2003[7]	Q1: Does the study population match the source population defined by the inclusion/exclusion criteria?"	Yes
	Q2: Was randomization adequately performed?	Yes
	Q3: Was allocation adequately concealed?	NR
	Q4: Did randomization result in equal practice and participant characteristics among study groups?	Yes
	Q5: Did the study maintain comparable groups?	NR
	Q6: In cases of high loss to follow-up (or differential loss to follow-up), is the impact assessed (e.g., through sensitivity analysis or other adjustment method)?	Yes
	Q7: Are outcomes assessed using valid and reliable measures?	Good
	Q8: Are outcomes implemented consistently across all study participants?	Good
	Q9: Were participants and intervention deliverers blinded during the study	NR
	Q10: Were the outcome assessors (e.g., data collectors) blinded to the intervention arm of participants?	Yes
	Q11: Was the intervention delivered to all members of the intervention group as described?	Fair
	Q12: Are the outcomes of interest pre-specified by the researchers?	Good
	Q13: Were the key confounding variables taken into account in the analysis and/or through randomization (e.g., successful randomization without loss to follow up, through matching, stratification, multivariate analysis, exclude based on confounding or other statistical adjustment)?	NR
	Q14: Are the statistical methods used to assess the primary benefit outcomes appropriate to the data?	Yes
	Q15: Are results believable taking study limitations into consideration?	Fair

Table C-2. Risk of bias for KQ 2 (continued)

Author, Year	Question	Response
Beaulieu, 2004[8]	Q1: Does the study population match the source population defined by the inclusion/exclusion criteria?"	No
	Q2: Was randomization adequately performed?	Yes
	Q3: Was allocation adequately concealed?	Yes
	Q4: Did randomization result in equal practice and participant characteristics among study groups?	Yes
	Q5: Did the study maintain comparable groups?	No
	Q6: In cases of high loss to follow-up (or differential loss to follow-up), is the impact assessed (e.g., through sensitivity analysis or other adjustment method)?	Yes
	Q7: Are outcomes assessed using valid and reliable measures?	Good
	Q8: Are outcomes implemented consistently across all study participants?	Good
	Q9: Were participants and intervention deliverers blinded during the study	NR
	Q10: Were the outcome assessors (e.g., data collectors) blinded to the intervention arm of participants?	NR
	Q11: Was the intervention delivered to all members of the intervention group as described?	NR
	Q12: Are the outcomes of interest pre-specified by the researchers?	Good
	Q13: Were the key confounding variables taken into account in the analysis and/or through randomization (e.g., successful randomization without loss to follow up, through matching, stratification, multivariate analysis, exclude based on confounding or other statistical adjustment)?	Yes
	Q14: Are the statistical methods used to assess the primary benefit outcomes appropriate to the data?	Yes
	Q15: Are results believable taking study limitations into consideration?	Fair
Becker, 2008[9]	Q1: Does the study population match the source population defined by the inclusion/exclusion criteria?"	Yes
	Q2: Was randomization adequately performed?	Yes
	Q3: Was allocation adequately concealed?	Yes
	Q4: Did randomization result in equal practice and participant characteristics among study groups?	Yes
	Q5: Did the study maintain comparable groups?	Yes
	Q6: In cases of high loss to follow-up (or differential loss to follow-up), is the impact assessed (e.g., through sensitivity analysis or other adjustment method)?	NA
	Q7: Are outcomes assessed using valid and reliable measures?	Good
	Q8: Are outcomes implemented consistently across all study participants?	Good
	Q9: Were participants and intervention deliverers blinded during the study	Fair
	Q10: Were the outcome assessors (e.g., data collectors) blinded to the intervention arm of participants?	NA
	Q11: Was the intervention delivered to all members of the intervention group as described?	Fair
	Q12: Are the outcomes of interest pre-specified by the researchers?	Good
	Q13: Were the key confounding variables taken into account in the analysis and/or through randomization (e.g., successful randomization without loss to follow up, through matching, stratification, multivariate analysis, exclude based on confounding or other statistical adjustment)?	Yes
	Q14: Are the statistical methods used to assess the primary benefit outcomes appropriate to the data?	Yes
	Q15: Are results believable taking study limitations into consideration?	Good

Table C-2. Risk of bias for KQ 2 (continued)

Author, Year	Question	Response
Bekkering, 2005[10]	Q1: Does the study population match the source population defined by the inclusion/exclusion criteria?"	Yes
	Q2: Was randomization adequately performed?	Yes
	Q3: Was allocation adequately concealed?	Yes
	Q4: Did randomization result in equal practice and participant characteristics among study groups?	Yes
	Q5: Did the study maintain comparable groups?	NR
	Q6: In cases of high loss to follow-up (or differential loss to follow-up), is the impact assessed (e.g., through sensitivity analysis or other adjustment method)?	Yes
	Q7: Are outcomes assessed using valid and reliable measures?	Good
	Q8: Are outcomes implemented consistently across all study participants?	Good
	Q9: Were participants and intervention deliverers blinded during the study	Good
	Q10: Were the outcome assessors (e.g., data collectors) blinded to the intervention arm of participants?	Yes
	Q11: Was the intervention delivered to all members of the intervention group as described?	Fair
	Q12: Are the outcomes of interest pre-specified by the researchers?	Good
	Q13: Were the key confounding variables taken into account in the analysis and/or through randomization (e.g., successful randomization without loss to follow up, through matching, stratification, multivariate analysis, exclude based on confounding or other statistical adjustment)?	Yes
	Q14: Are the statistical methods used to assess the primary benefit outcomes appropriate to the data?	Yes
	Q15: Are results believable taking study limitations into consideration?	Good
Bekkering, 2005[11]	Q1: Does the study population match the source population defined by the inclusion/exclusion criteria?"	Fair
	Q2: Was randomization adequately performed?	Yes
	Q3: Was allocation adequately concealed?	Yes
	Q4: Did randomization result in equal practice and participant characteristics among study groups?	Yes
	Q5: Did the study maintain comparable groups?	Yes
	Q6: In cases of high loss to follow-up (or differential loss to follow-up), is the impact assessed (e.g., through sensitivity analysis or other adjustment method)?	NA
	Q7: Are outcomes assessed using valid and reliable measures?	Good
	Q8: Are outcomes implemented consistently across all study participants?	Good
	Q9: Were participants and intervention deliverers blinded during the study	Fair
	Q10: Were the outcome assessors (e.g., data collectors) blinded to the intervention arm of participants?	NR
	Q11: Was the intervention delivered to all members of the intervention group as described?	Good
	Q12: Are the outcomes of interest pre-specified by the researchers?	Fair
	Q13: Were the key confounding variables taken into account in the analysis and/or through randomization (e.g., successful randomization without loss to follow up, through matching, stratification, multivariate analysis, exclude based on confounding or other statistical adjustment)?	Yes
	Q14: Are the statistical methods used to assess the primary benefit outcomes appropriate to the data?	Yes
	Q15: Are results believable taking study limitations into consideration?	Good

Table C-2. Risk of bias for KQ 2 (continued)

Author, Year	Question	Response
Bishop, 2006[12]	Q1: Does the study population match the source population defined by the inclusion/exclusion criteria?"	Yes
	Q2: Was randomization adequately performed?	Yes
	Q3: Was allocation adequately concealed?	NR
	Q4: Did randomization result in equal practice and participant characteristics among study groups?	Yes
	Q5: Did the study maintain comparable groups?	Yes
	Q6: In cases of high loss to follow-up (or differential loss to follow-up), is the impact assessed (e.g., through sensitivity analysis or other adjustment method)?	NA
	Q7: Are outcomes assessed using valid and reliable measures?	Good
	Q8: Are outcomes implemented consistently across all study participants?	Good
	Q9: Were participants and intervention deliverers blinded during the study	Fair
	Q10: Were the outcome assessors (e.g., data collectors) blinded to the intervention arm of participants?	NR
	Q11: Was the intervention delivered to all members of the intervention group as described?	NR
	Q12: Are the outcomes of interest pre-specified by the researchers?	Good
	Q13: Were the key confounding variables taken into account in the analysis and/or through randomization (e.g., successful randomization without loss to follow up, through matching, stratification, multivariate analysis, exclude based on confounding or other statistical adjustment)?	No
	Q14: Are the statistical methods used to assess the primary benefit outcomes appropriate to the data?	Yes
	Q15: Are results believable taking study limitations into consideration?	Good
Campbell, 2004[13]	Q1: Does the study population match the source population defined by the inclusion/exclusion criteria?"	No
	Q2: Was randomization adequately performed?	No
	Q3: Was allocation adequately concealed?	Yes
	Q4: Did randomization result in equal practice and participant characteristics among study groups?	Yes
	Q5: Did the study maintain comparable groups?	No
	Q6: In cases of high loss to follow-up (or differential loss to follow-up), is the impact assessed (e.g., through sensitivity analysis or other adjustment method)?	No
	Q7: Are outcomes assessed using valid and reliable measures?	Good
	Q8: Are outcomes implemented consistently across all study participants?	Fair
	Q9: Were participants and intervention deliverers blinded during the study	Poor
	Q10: Were the outcome assessors (e.g., data collectors) blinded to the intervention arm of participants?	NA
	Q11: Was the intervention delivered to all members of the intervention group as described?	Fair
	Q12: Are the outcomes of interest pre-specified by the researchers?	Good
	Q13: Were the key confounding variables taken into account in the analysis and/or through randomization (e.g., successful randomization without loss to follow up, through matching, stratification, multivariate analysis, exclude based on confounding or other statistical adjustment)?	Yes
	Q14: Are the statistical methods used to assess the primary benefit outcomes appropriate to the data?	Yes
	Q15: Are results believable taking study limitations into consideration?	Fair

Table C-2. Risk of bias for KQ 2 (continued)

Author, Year	Question	Response
Carney, 2005[14]	Q1: Does the study population match the source population defined by the inclusion/exclusion criteria?"	Yes
	Q2: Was randomization adequately performed?	NR
	Q3: Was allocation adequately concealed?	NR
	Q4: Did randomization result in equal practice and participant characteristics among study groups?	Yes
	Q5: Did the study maintain comparable groups?	Yes
	Q6: In cases of high loss to follow-up (or differential loss to follow-up), is the impact assessed (e.g., through sensitivity analysis or other adjustment method)?	NA
	Q7: Are outcomes assessed using valid and reliable measures?	Fair
	Q8: Are outcomes implemented consistently across all study participants?	Good
	Q9: Were participants and intervention deliverers blinded during the study	Fair
	Q10: Were the outcome assessors (e.g., data collectors) blinded to the intervention arm of participants?	NR
	Q11: Was the intervention delivered to all members of the intervention group as described?	Good
	Q12: Are the outcomes of interest pre-specified by the researchers?	Good
	Q13: Were the key confounding variables taken into account in the analysis and/or through randomization (e.g., successful randomization without loss to follow up, through matching, stratification, multivariate analysis, exclude based on confounding or other statistical adjustment)?	Yes
	Q14: Are the statistical methods used to assess the primary benefit outcomes appropriate to the data?	Yes
	Q15: Are results believable taking study limitations into consideration?	Fair
Christakis, 2006[15]	Q1: Does the study population match the source population defined by the inclusion/exclusion criteria?"	No
	Q2: Was randomization adequately performed?	Yes
	Q3: Was allocation adequately concealed?	Yes
	Q4: Did randomization result in equal practice and participant characteristics among study groups?	Yes
	Q5: Did the study maintain comparable groups?	Yes
	Q6: In cases of high loss to follow-up (or differential loss to follow-up), is the impact assessed (e.g., through sensitivity analysis or other adjustment method)?	NA
	Q7: Are outcomes assessed using valid and reliable measures?	Fair
	Q8: Are outcomes implemented consistently across all study participants?	Good
	Q9: Were participants and intervention deliverers blinded during the study	Fair
	Q10: Were the outcome assessors (e.g., data collectors) blinded to the intervention arm of participants?	Yes
	Q11: Was the intervention delivered to all members of the intervention group as described?	Fair
	Q12: Are the outcomes of interest pre-specified by the researchers?	Good
	Q13: Were the key confounding variables taken into account in the analysis and/or through randomization (e.g., successful randomization without loss to follow up, through matching, stratification, multivariate analysis, exclude based on confounding or other statistical adjustment)?	Yes
	Q14: Are the statistical methods used to assess the primary benefit outcomes appropriate to the data?	Yes
	Q15: Are results believable taking study limitations into consideration?	Fair

Table C-2. Risk of bias for KQ 2 (continued)

Author, Year	Question	Response
Church, 2004[16]	Q1: Does the study population match the source population defined by the inclusion/exclusion criteria?"	Yes
	Q2: Was randomization adequately performed?	NR
	Q3: Was allocation adequately concealed?	NR
	Q4: Did randomization result in equal practice and participant characteristics among study groups?	NR
	Q5: Did the study maintain comparable groups?	Yes
	Q6: In cases of high loss to follow-up (or differential loss to follow-up), is the impact assessed (e.g., through sensitivity analysis or other adjustment method)?	NA
	Q7: Are outcomes assessed using valid and reliable measures?	Fair
	Q8: Are outcomes implemented consistently across all study participants?	Poor
	Q9: Were participants and intervention deliverers blinded during the study	NR
	Q10: Were the outcome assessors (e.g., data collectors) blinded to the intervention arm of participants?	No
	Q11: Was the intervention delivered to all members of the intervention group as described?	Poor
	Q12: Are the outcomes of interest pre-specified by the researchers?	Good
	Q13: Were the key confounding variables taken into account in the analysis and/or through randomization (e.g., successful randomization without loss to follow up, through matching, stratification, multivariate analysis, exclude based on confounding or other statistical adjustment)?	No
	Q14: Are the statistical methods used to assess the primary benefit outcomes appropriate to the data?	No
	Q15: Are results believable taking study limitations into consideration?	Poor
Daniels, 2005[17]	Q1: Does the study population match the source population defined by the inclusion/exclusion criteria?"	Poor
	Q2: Was randomization adequately performed?	NR
	Q3: Was allocation adequately concealed?	NR
	Q4: Did randomization result in equal practice and participant characteristics among study groups?	No
	Q5: Did the study maintain comparable groups?	NR
	Q6: In cases of high loss to follow-up (or differential loss to follow-up), is the impact assessed (e.g., through sensitivity analysis or other adjustment method)?	NR
	Q7: Are outcomes assessed using valid and reliable measures?	NR
	Q8: Are outcomes implemented consistently across all study participants?	Good
	Q9: Were participants and intervention deliverers blinded during the study	NR
	Q10: Were the outcome assessors (e.g., data collectors) blinded to the intervention arm of participants?	NR
	Q11: Was the intervention delivered to all members of the intervention group as described?	Poor
	Q12: Are the outcomes of interest pre-specified by the researchers?	Good
	Q13: Were the key confounding variables taken into account in the analysis and/or through randomization (e.g., successful randomization without loss to follow up, through matching, stratification, multivariate analysis, exclude based on confounding or other statistical adjustment)?	No
	Q14: Are the statistical methods used to assess the primary benefit outcomes appropriate to the data?	No
	Q15: Are results believable taking study limitations into consideration?	Poor

Table C-2. Risk of bias for KQ 2 (continued)

Author, Year	Question	Response
Davis, 2004[18]	Q1: Does the study population match the source population defined by the inclusion/exclusion criteria?"	Yes
	Q2: Was randomization adequately performed?	Yes
	Q3: Was allocation adequately concealed?	Yes
	Q4: Did randomization result in equal practice and participant characteristics among study groups?	Yes
	Q5: Did the study maintain comparable groups?	No
	Q6: In cases of high loss to follow-up (or differential loss to follow-up), is the impact assessed (e.g., through sensitivity analysis or other adjustment method)?	Yes
	Q7: Are outcomes assessed using valid and reliable measures?	Good
	Q8: Are outcomes implemented consistently across all study participants?	Good
	Q9: Were participants and intervention deliverers blinded during the study	NR
	Q10: Were the outcome assessors (e.g., data collectors) blinded to the intervention arm of participants?	NR
	Q11: Was the intervention delivered to all members of the intervention group as described?	Poor
	Q12: Are the outcomes of interest pre-specified by the researchers?	Good
	Q13: Were the key confounding variables taken into account in the analysis and/or through randomization (e.g., successful randomization without loss to follow up, through matching, stratification, multivariate analysis, exclude based on confounding or other statistical adjustment)?	Yes
	Q14: Are the statistical methods used to assess the primary benefit outcomes appropriate to the data?	Yes
	Q15: Are results believable taking study limitations into consideration?	Fair
Eaton, 2011[19]	Q1: Does the study population match the source population defined by the inclusion/exclusion criteria?"	Fair
	Q2: Was randomization adequately performed?	Yes
	Q3: Was allocation adequately concealed?	NR
	Q4: Did randomization result in equal practice and participant characteristics among study groups?	Yes
	Q5: Did the study maintain comparable groups?	NR
	Q6: In cases of high loss to follow-up (or differential loss to follow-up), is the impact assessed (e.g., through sensitivity analysis or other adjustment method)?	Yes
	Q7: Are outcomes assessed using valid and reliable measures?	Good
	Q8: Are outcomes implemented consistently across all study participants?	NR
	Q9: Were participants and intervention deliverers blinded during the study	Poor
	Q10: Were the outcome assessors (e.g., data collectors) blinded to the intervention arm of participants?	Yes
	Q11: Was the intervention delivered to all members of the intervention group as described?	NR
	Q12: Are the outcomes of interest pre-specified by the researchers?	Good
	Q13: Were the key confounding variables taken into account in the analysis and/or through randomization (e.g., successful randomization without loss to follow up, through matching, stratification, multivariate analysis, exclude based on confounding or other statistical adjustment)?	Yes
	Q14: Are the statistical methods used to assess the primary benefit outcomes appropriate to the data?	Yes
	Q15: Are results believable taking study limitations into consideration?	Good

Table C-2. Risk of bias for KQ 2 (continued)

Author, Year	Question	Response
Elder, 2005[2]	Q1: Does the study population match the source population defined by the inclusion/exclusion criteria?"	Yes
	Q2: Was randomization adequately performed?	NR
	Q3: Was allocation adequately concealed?	NR
	Q4: Did randomization result in equal practice and participant characteristics among study groups?	No
	Q5: Did the study maintain comparable groups?	Yes
	Q6: In cases of high loss to follow-up (or differential loss to follow-up), is the impact assessed (e.g., through sensitivity analysis or other adjustment method)?	NA
	Q7: Are outcomes assessed using valid and reliable measures?	Good
	Q8: Are outcomes implemented consistently across all study participants?	Good
	Q9: Were participants and intervention deliverers blinded during the study?	Fair
	Q10: Were the outcome assessors (e.g., data collectors) blinded to the intervention arm of participants?	Yes
	Q11: Was the intervention delivered to all members of the intervention group as described?	Good
	Q12: Are the outcomes of interest pre-specified by the researchers?	Good
	Q13: Were the key confounding variables taken into account in the analysis and/or through randomization (e.g., successful randomization without loss to follow up, through matching, stratification, multivariate analysis, exclude based on confounding or other statistical adjustment)?	Yes
	Q14: Are the statistical methods used to assess the primary benefit outcomes appropriate to the data?	Yes
	Q15: Are results believable taking study limitations into consideration?	Fair
Feldstein, 2006[20]	Q1: Does the study population match the source population defined by the inclusion/exclusion criteria?"	Good
	Q2: Was randomization adequately performed?	Yes
	Q3: Was allocation adequately concealed?	No
	Q4: Did randomization result in equal practice and participant characteristics among study groups?	Yes
	Q5: Did the study maintain comparable groups?	Yes
	Q6: In cases of high loss to follow-up (or differential loss to follow-up), is the impact assessed (e.g., through sensitivity analysis or other adjustment method)?	NA
	Q7: Are outcomes assessed using valid and reliable measures?	Good
	Q8: Are outcomes implemented consistently across all study participants?	Good
	Q9: Were participants and intervention deliverers blinded during the study	Poor
	Q10: Were the outcome assessors (e.g., data collectors) blinded to the intervention arm of participants?	Yes
	Q11: Was the intervention delivered to all members of the intervention group as described?	Good
	Q12: Are the outcomes of interest pre-specified by the researchers?	Good
	Q13: Were the key confounding variables taken into account in the analysis and/or through randomization (e.g., successful randomization without loss to follow up, through matching, stratification, multivariate analysis, exclude based on confounding or other statistical adjustment)?	Yes
	Q14: Are the statistical methods used to assess the primary benefit outcomes appropriate to the data?	Yes
	Q15: Are results believable taking study limitations into consideration?	Good

Table C-2. Risk of bias for KQ 2

Author, Year	Question	Response
Figuerias, [21]	Q1: Does the study population match the source population defined by the inclusion/exclusion criteria?"	Poor
	Q2: Was randomization adequately performed?	NR
	Q3: Was allocation adequately concealed?	NR
	Q4: Did randomization result in equal practice and participant characteristics among study groups?	No
	Q5: Did the study maintain comparable groups?	No
	Q6: In cases of high loss to follow-up (or differential loss to follow-up), is the impact assessed (e.g., through sensitivity analysis or other adjustment method)?	Yes
	Q7: Are outcomes assessed using valid and reliable measures?	Yes
	Q8: Are outcomes implemented consistently across all study participants?	Yes
	Q9: Were participants and intervention deliverers blinded during the study	NR
	Q10: Were the outcome assessors (e.g., data collectors) blinded to the intervention arm of participants?	NR
	Q11: Was the intervention delivered to all members of the intervention group as described?	Fair
	Q12: Are the outcomes of interest pre-specified by the researchers?	Good
	Q13: Were the key confounding variables taken into account in the analysis and/or through randomization (e.g., successful randomization without loss to follow up, through matching, stratification, multivariate analysis, exclude based on confounding or other statistical adjustment)?	Yes
	Q14: Are the statistical methods used to assess the primary benefit outcomes appropriate to the data?	Yes
	Q15: Are results believable taking study limitations into consideration?	Poor
Gatellari, 2005[22]	Q1: Does the study population match the source population defined by the inclusion/exclusion criteria?"	Fair
	Q2: Was randomization adequately performed?	Yes
	Q3: Was allocation adequately concealed?	Yes
	Q4: Did randomization result in equal practice and participant characteristics among study groups?	Yes
	Q5: Did the study maintain comparable groups?	Yes
	Q6: In cases of high loss to follow-up (or differential loss to follow-up), is the impact assessed (e.g., through sensitivity analysis or other adjustment method)?	NA
	Q7: Are outcomes assessed using valid and reliable measures?	Fair
	Q8: Are outcomes implemented consistently across all study participants?	Fair
	Q9: Were participants and intervention deliverers blinded during the study	Good
	Q10: Were the outcome assessors (e.g., data collectors) blinded to the intervention arm of participants?	No
	Q11: Was the intervention delivered to all members of the intervention group as described?	Good
	Q12: Are the outcomes of interest pre-specified by the researchers?	Good
	Q13: Were the key confounding variables taken into account in the analysis and/or through randomization (e.g., successful randomization without loss to follow up, through matching, stratification, multivariate analysis, exclude based on confounding or other statistical adjustment)?	Yes
	Q14: Are the statistical methods used to assess the primary benefit outcomes appropriate to the data?	Yes
	Q15: Are results believable taking study limitations into consideration?	Fair

Table C-2. Risk of bias for KQ 2

Author, Year	Question	Response
Hagmolen, 2008[23]	Q1: Does the study population match the source population defined by the inclusion/exclusion criteria?"	Yes
	Q2: Was randomization adequately performed?	NR
	Q3: Was allocation adequately concealed?	Yes
	Q4: Did randomization result in equal practice and participant characteristics among study groups?	Yes
	Q5: Did the study maintain comparable groups?	Yes
	Q6: In cases of high loss to follow-up (or differential loss to follow-up), is the impact assessed (e.g., through sensitivity analysis or other adjustment method)?	NA
	Q7: Are outcomes assessed using valid and reliable measures?	Good
	Q8: Are outcomes implemented consistently across all study participants?	NR
	Q9: Were participants and intervention deliverers blinded during the study	Fair
	Q10: Were the outcome assessors (e.g., data collectors) blinded to the intervention arm of participants?	No
	Q11: Was the intervention delivered to all members of the intervention group as described?	Poor
	Q12: Are the outcomes of interest pre-specified by the researchers?	Good
	Q13: Were the key confounding variables taken into account in the analysis and/or through randomization (e.g., successful randomization without loss to follow up, through matching, stratification, multivariate analysis, exclude based on confounding or other statistical adjustment)?	Yes
	Q14: Are the statistical methods used to assess the primary benefit outcomes appropriate to the data?	Yes
	Q15: Are results believable taking study limitations into consideration?	Fair
Jain, 2006[24]	Q1: Does the study population match the source population defined by the inclusion/exclusion criteria?"	Yes
	Q2: Was randomization adequately performed?	Yes
	Q3: Was allocation adequately concealed?	No
	Q4: Did randomization result in equal practice and participant characteristics among study groups?	Yes
	Q5: Did the study maintain comparable groups?	Yes
	Q6: In cases of high loss to follow-up (or differential loss to follow-up), is the impact assessed (e.g., through sensitivity analysis or other adjustment method)?	NA
	Q7: Are outcomes assessed using valid and reliable measures?	Good
	Q8: Are outcomes implemented consistently across all study participants?	NR
	Q9: Were participants and intervention deliverers blinded during the study	Fair
	Q10: Were the outcome assessors (e.g., data collectors) blinded to the intervention arm of participants?	No
	Q11: Was the intervention delivered to all members of the intervention group as described?	Poor
	Q12: Are the outcomes of interest pre-specified by the researchers?	Good
	Q13: Were the key confounding variables taken into account in the analysis and/or through randomization (e.g., successful randomization without loss to follow up, through matching, stratification, multivariate analysis, exclude based on confounding or other statistical adjustment)?	Good
	Q14: Are the statistical methods used to assess the primary benefit outcomes appropriate to the data?	Yes
	Q15: Are results believable taking study limitations into consideration?	Fair

Table C-2. Risk of bias for KQ 2 (continued)

Author, Year	Question	Response
Jousimaa, 2002[25]	Q1: Does the study population match the source population defined by the inclusion/exclusion criteria?"	Fair
	Q2: Was randomization adequately performed?	Yes
	Q3: Was allocation adequately concealed?	NR
	Q4: Did randomization result in equal practice and participant characteristics among study groups?	Yes
	Q5: Did the study maintain comparable groups?	Yes
	Q6: In cases of high loss to follow-up (or differential loss to follow-up), is the impact assessed (e.g., through sensitivity analysis or other adjustment method)?	NA
	Q7: Are outcomes assessed using valid and reliable measures?	Good
	Q8: Are outcomes implemented consistently across all study participants?	Good
	Q9: Were participants and intervention deliverers blinded during the study	NR
	Q10: Were the outcome assessors (e.g., data collectors) blinded to the intervention arm of participants?	Yes
	Q11: Was the intervention delivered to all members of the intervention group as described?	Good
	Q12: Are the outcomes of interest pre-specified by the researchers?	Good
	Q13: Were the key confounding variables taken into account in the analysis and/or through randomization (e.g., successful randomization without loss to follow up, through matching, stratification, multivariate analysis, exclude based on confounding or other statistical adjustment)?	Yes
	Q14: Are the statistical methods used to assess the primary benefit outcomes appropriate to the data?	Yes
	Q15: Are results believable taking study limitations into consideration?	Good
Junghaus, 2007[26]	Q1: Does the study population match the source population defined by the inclusion/exclusion criteria?"	Yes
	Q2: Was randomization adequately performed?	Yes
	Q3: Was allocation adequately concealed?	NR
	Q4: Did randomization result in equal practice and participant characteristics among study groups?	Yes
	Q5: Did the study maintain comparable groups?	Yes
	Q6: In cases of high loss to follow-up (or differential loss to follow-up), is the impact assessed (e.g., through sensitivity analysis or other adjustment method)?	NA
	Q7: Are outcomes assessed using valid and reliable measures?	Good
	Q8: Are outcomes implemented consistently across all study participants?	NR
	Q9: Were participants and intervention deliverers blinded during the study	Fair
	Q10: Were the outcome assessors (e.g., data collectors) blinded to the intervention arm of participants?	No
	Q11: Was the intervention delivered to all members of the intervention group as described?	NR
	Q12: Are the outcomes of interest pre-specified by the researchers?	Good
	Q13: Were the key confounding variables taken into account in the analysis and/or through randomization (e.g., successful randomization without loss to follow up, through matching, stratification, multivariate analysis, exclude based on confounding or other statistical adjustment)?	Yes
	Q14: Are the statistical methods used to assess the primary benefit outcomes appropriate to the data?	Yes
	Q15: Are results believable taking study limitations into consideration?	Fair

Table C-2. Risk of bias for KQ 2 (continued)

Author, Year	Question	Response
Kennedy, 2003[27]	Q1: Does the study population match the source population defined by the inclusion/exclusion criteria?"	No
	Q2: Was randomization adequately performed?	Yes
	Q3: Was allocation adequately concealed?	Yes
	Q4: Did randomization result in equal practice and participant characteristics among study groups?	Yes
	Q5: Did the study maintain comparable groups?	No
	Q6: In cases of high loss to follow-up (or differential loss to follow-up), is the impact assessed (e.g., through sensitivity analysis or other adjustment method)?	Yes
	Q7: Are outcomes assessed using valid and reliable measures?	Good
	Q8: Are outcomes implemented consistently across all study participants?	Good
	Q9: Were participants and intervention deliverers blinded during the study	Fair
	Q10: Were the outcome assessors (e.g., data collectors) blinded to the intervention arm of participants?	No
	Q11: Was the intervention delivered to all members of the intervention group as described?	Fair
	Q12: Are the outcomes of interest pre-specified by the researchers?	Good
	Q13: Were the key confounding variables taken into account in the analysis and/or through randomization (e.g., successful randomization without loss to follow up, through matching, stratification, multivariate analysis, exclude based on confounding or other statistical adjustment)?	Yes
	Q14: Are the statistical methods used to assess the primary benefit outcomes appropriate to the data?	Yes
	Q15: Are results believable taking study limitations into consideration?	Good
King, 2007[28]	Q1: Does the study population match the source population defined by the inclusion/exclusion criteria?"	Fair
	Q2: Was randomization adequately performed?	Yes
	Q3: Was allocation adequately concealed?	Yes
	Q4: Did randomization result in equal practice and participant characteristics among study groups?	Yes
	Q5: Did the study maintain comparable groups?	Yes
	Q6: In cases of high loss to follow-up (or differential loss to follow-up), is the impact assessed (e.g., through sensitivity analysis or other adjustment method)?	Yes
	Q7: Are outcomes assessed using valid and reliable measures?	Good
	Q8: Are outcomes implemented consistently across all study participants?	Good
	Q9: Were participants and intervention deliverers blinded during the study	NR
	Q10: Were the outcome assessors (e.g., data collectors) blinded to the intervention arm of participants?	Yes
	Q11: Was the intervention delivered to all members of the intervention group as described?	Good
	Q12: Are the outcomes of interest pre-specified by the researchers?	Good
	Q13: Were the key confounding variables taken into account in the analysis and/or through randomization (e.g., successful randomization without loss to follow up, through matching, stratification, multivariate analysis, exclude based on confounding or other statistical adjustment)?	Yes
	Q14: Are the statistical methods used to assess the primary benefit outcomes appropriate to the data?	Yes
	Q15: Are results believable taking study limitations into consideration?	Good

Table C-2. Risk of bias for KQ 2 (continued)

Author, Year	Question	Response
Laprise, 2009[29]	Q1: Does the study population match the source population defined by the inclusion/exclusion criteria?"	Yes
	Q2: Was randomization adequately performed?	Yes
	Q3: Was allocation adequately concealed?	NR
	Q4: Did randomization result in equal practice and participant characteristics among study groups?	Yes
	Q5: Did the study maintain comparable groups?	NR
	Q6: In cases of high loss to follow-up (or differential loss to follow-up), is the impact assessed (e.g., through sensitivity analysis or other adjustment method)?	NR
	Q7: Are outcomes assessed using valid and reliable measures?	Good
	Q8: Are outcomes implemented consistently across all study participants?	NR
	Q9: Were participants and intervention deliverers blinded during the study	Poor
	Q10: Were the outcome assessors (e.g., data collectors) blinded to the intervention arm of participants?	No
	Q11: Was the intervention delivered to all members of the intervention group as described?	NR
	Q12: Are the outcomes of interest pre-specified by the researchers?	Good
	Q13: Were the key confounding variables taken into account in the analysis and/or through randomization (e.g., successful randomization without loss to follow up, through matching, stratification, multivariate analysis, exclude based on confounding or other statistical adjustment)?	Yes
	Q14: Are the statistical methods used to assess the primary benefit outcomes appropriate to the data?	Yes
	Q15: Are results believable taking study limitations into consideration?	Fair
Lien, 2007[30]	Q1: Does the study population match the source population defined by the inclusion/exclusion criteria?"	Fair
	Q2: Was randomization adequately performed?	Yes
	Q3: Was allocation adequately concealed?	Yes
	Q4: Did randomization result in equal practice and participant characteristics among study groups?	Yes
	Q5: Did the study maintain comparable groups?	Yes
	Q6: In cases of high loss to follow-up (or differential loss to follow-up), is the impact assessed (e.g., through sensitivity analysis or other adjustment method)?	Yes
	Q7: Are outcomes assessed using valid and reliable measures?	Good
	Q8: Are outcomes implemented consistently across all study participants?	Good
	Q9: Were participants and intervention deliverers blinded during the study	Fair
	Q10: Were the outcome assessors (e.g., data collectors) blinded to the intervention arm of participants?	Yes
	Q11: Was the intervention delivered to all members of the intervention group as described?	Fair
	Q12: Are the outcomes of interest pre-specified by the researchers?	Good
	Q13: Were the key confounding variables taken into account in the analysis and/or through randomization (e.g., successful randomization without loss to follow up, through matching, stratification, multivariate analysis, exclude based on confounding or other statistical adjustment)?	Yes
	Q14: Are the statistical methods used to assess the primary benefit outcomes appropriate to the data?	Yes
	Q15: Are results believable taking study limitations into consideration?	Good

Table C-2. Risk of bias for KQ 2 (continued)

Author, Year	Question	Response
Manfredi, 2011[31]	Q1: Does the study population match the source population defined by the inclusion/exclusion criteria?"	Fair
	Q2: Was randomization adequately performed?	NR
	Q3: Was allocation adequately concealed?	NR
	Q4: Did randomization result in equal practice and participant characteristics among study groups?	NR
	Q5: Did the study maintain comparable groups?	No
	Q6: In cases of high loss to follow-up (or differential loss to follow-up), is the impact assessed (e.g., through sensitivity analysis or other adjustment method)?	NR
	Q7: Are outcomes assessed using valid and reliable measures?	NR
	Q8: Are outcomes implemented consistently across all study participants?	NR
	Q9: Were participants and intervention deliverers blinded during the study	NR
	Q10: Were the outcome assessors (e.g., data collectors) blinded to the intervention arm of participants?	NR
	Q11: Was the intervention delivered to all members of the intervention group as described?	Poor
	Q12: Are the outcomes of interest pre-specified by the researchers?	Good
	Q13: Were the key confounding variables taken into account in the analysis and/or through randomization (e.g., successful randomization without loss to follow up, through matching, stratification, multivariate analysis, exclude based on confounding or other statistical adjustment)?	Yes
	Q14: Are the statistical methods used to assess the primary benefit outcomes appropriate to the data?	Yes
	Q15: Are results believable taking study limitations into consideration?	Poor
Marcus, 2007[32]	Q1: Does the study population match the source population defined by the inclusion/exclusion criteria?"	Poor
	Q2: Was randomization adequately performed?	NR
	Q3: Was allocation adequately concealed?	NR
	Q4: Did randomization result in equal practice and participant characteristics among study groups?	No
	Q5: Did the study maintain comparable groups?	Yes
	Q6: In cases of high loss to follow-up (or differential loss to follow-up), is the impact assessed (e.g., through sensitivity analysis or other adjustment method)?	Yes
	Q7: Are outcomes assessed using valid and reliable measures?	Good
	Q8: Are outcomes implemented consistently across all study participants?	Good
	Q9: Were participants and intervention deliverers blinded during the study	NR
	Q10: Were the outcome assessors (e.g., data collectors) blinded to the intervention arm of participants?	NR
	Q11: Was the intervention delivered to all members of the intervention group as described?	NR
	Q12: Are the outcomes of interest pre-specified by the researchers?	Good
	Q13: Were the key confounding variables taken into account in the analysis and/or through randomization (e.g., successful randomization without loss to follow up, through matching, stratification, multivariate analysis, exclude based on confounding or other statistical adjustment)?	Yes
	Q14: Are the statistical methods used to assess the primary benefit outcomes appropriate to the data?	Yes
	Q15: Are results believable taking study limitations into consideration?	Fair

Table C-2. Risk of bias for KQ 2 (continued)

Author, Year	Question	Response
Maxwell, 2010[33]	Q1: Does the study population match the source population defined by the inclusion/exclusion criteria?"	Fair
	Q2: Was randomization adequately performed?	NR
	Q3: Was allocation adequately concealed?	NR
	Q4: Did randomization result in equal practice and participant characteristics among study groups?	Yes
	Q5: Did the study maintain comparable groups?	Yes
	Q6: In cases of high loss to follow-up (or differential loss to follow-up), is the impact assessed (e.g., through sensitivity analysis or other adjustment method)?	Yes
	Q7: Are outcomes assessed using valid and reliable measures?	Fair
	Q8: Are outcomes implemented consistently across all study participants?	Fair
	Q9: Were participants and intervention deliverers blinded during the study	No
	Q10: Were the outcome assessors (e.g., data collectors) blinded to the intervention arm of participants?	NR
	Q11: Was the intervention delivered to all members of the intervention group as described?	Good
	Q12: Are the outcomes of interest pre-specified by the researchers?	Good
	Q13: Were the key confounding variables taken into account in the analysis and/or through randomization (e.g., successful randomization without loss to follow up, through matching, stratification, multivariate analysis, exclude based on confounding or other statistical adjustment)?	Yes
	Q14: Are the statistical methods used to assess the primary benefit outcomes appropriate to the data?	Yes
	Q15: Are results believable taking study limitations into consideration?	Fair
Murtaugh, 2005[34]	Q1: Does the study population match the source population defined by the inclusion/exclusion criteria?"	NR
	Q2: Was randomization adequately performed?	NR
	Q3: Was allocation adequately concealed?	NR
	Q4: Did randomization result in equal practice and participant characteristics among study groups?	No
	Q5: Did the study maintain comparable groups?	NR
	Q6: In cases of high loss to follow-up (or differential loss to follow-up), is the impact assessed (e.g., through sensitivity analysis or other adjustment method)?	NA
	Q7: Are outcomes assessed using valid and reliable measures?	Fair
	Q8: Are outcomes implemented consistently across all study participants?	NR
	Q9: Were participants and intervention deliverers blinded during the study	NR
	Q10: Were the outcome assessors (e.g., data collectors) blinded to the intervention arm of participants?	Yes
	Q11: Was the intervention delivered to all members of the intervention group as described?	NR
	Q12: Are the outcomes of interest pre-specified by the researchers?	Good
	Q13: Were the key confounding variables taken into account in the analysis and/or through randomization (e.g., successful randomization without loss to follow up, through matching, stratification, multivariate analysis, exclude based on confounding or other statistical adjustment)?	Yes
	Q14: Are the statistical methods used to assess the primary benefit outcomes appropriate to the data?	No
	Q15: Are results believable taking study limitations into consideration?	Fair

Table C-2. Risk of bias for KQ 2 (continued)

Author, Year	Question	Response
Paradis, 2011[35]	Q1: Does the study population match the source population defined by the inclusion/exclusion criteria?"	Poor
	Q2: Was randomization adequately performed?	No
	Q3: Was allocation adequately concealed?	NR
	Q4: Did randomization result in equal practice and participant characteristics among study groups?	No
	Q5: Did the study maintain comparable groups?	Yes
	Q6: In cases of high loss to follow-up (or differential loss to follow-up), is the impact assessed (e.g., through sensitivity analysis or other adjustment method)?	NA
	Q7: Are outcomes assessed using valid and reliable measures?	Fair
	Q8: Are outcomes implemented consistently across all study participants?	Good
	Q9: Were participants and intervention deliverers blinded during the study	NR
	Q10: Were the outcome assessors (e.g., data collectors) blinded to the intervention arm of participants?	Yes
	Q11: Was the intervention delivered to all members of the intervention group as described?	Good
	Q12: Are the outcomes of interest pre-specified by the researchers?	Good
	Q13: Were the key confounding variables taken into account in the analysis and/or through randomization (e.g., successful randomization without loss to follow up, through matching, stratification, multivariate analysis, exclude based on confounding or other statistical adjustment)?	Yes
	Q14: Are the statistical methods used to assess the primary benefit outcomes appropriate to the data?	Yes
	Q15: Are results believable taking study limitations into consideration?	Fair
Partin, 2004[36]	Q1: Does the study population match the source population defined by the inclusion/exclusion criteria?"	Fair
	Q2: Was randomization adequately performed?	Yes
	Q3: Was allocation adequately concealed?	Yes
	Q4: Did randomization result in equal practice and participant characteristics among study groups?	Yes
	Q5: Did the study maintain comparable groups?	Yes
	Q6: In cases of high loss to follow-up (or differential loss to follow-up), is the impact assessed (e.g., through sensitivity analysis or other adjustment method)?	NA
	Q7: Are outcomes assessed using valid and reliable measures?	Fair
	Q8: Are outcomes implemented consistently across all study participants?	NR
	Q9: Were participants and intervention deliverers blinded during the study	Good
	Q10: Were the outcome assessors (e.g., data collectors) blinded to the intervention arm of participants?	NA
	Q11: Was the intervention delivered to all members of the intervention group as described?	Poor
	Q12: Are the outcomes of interest pre-specified by the researchers?	Fair
	Q13: Were the key confounding variables taken into account in the analysis and/or through randomization (e.g., successful randomization without loss to follow up, through matching, stratification, multivariate analysis, exclude based on confounding or other statistical adjustment)?	Yes
	Q14: Are the statistical methods used to assess the primary benefit outcomes appropriate to the data?	Yes
	Q15: Are results believable taking study limitations into consideration?	Fair

Table C-2. Risk of bias for KQ 2 (continued)

Author, Year	Question	Response
Rahme, 2005[37]	Q1: Does the study population match the source population defined by the inclusion/exclusion criteria?"	Yes
	Q2: Was randomization adequately performed?	NR
	Q3: Was allocation adequately concealed?	NR
	Q4: Did randomization result in equal practice and participant characteristics among study groups?	Yes
	Q5: Did the study maintain comparable groups?	NR
	Q6: In cases of high loss to follow-up (or differential loss to follow-up), is the impact assessed (e.g., through sensitivity analysis or other adjustment method)?	NR
	Q7: Are outcomes assessed using valid and reliable measures?	Good
	Q8: Are outcomes implemented consistently across all study participants?	Good
	Q9: Were participants and intervention deliverers blinded during the study	NR
	Q10: Were the outcome assessors (e.g., data collectors) blinded to the intervention arm of participants?	NR
	Q11: Was the intervention delivered to all members of the intervention group as described?	Poor
	Q12: Are the outcomes of interest pre-specified by the researchers?	Good
	Q13: Were the key confounding variables taken into account in the analysis and/or through randomization (e.g., successful randomization without loss to follow up, through matching, stratification, multivariate analysis, exclude based on confounding or other statistical adjustment)?	Yes
	Q14: Are the statistical methods used to assess the primary benefit outcomes appropriate to the data?	Yes
	Q15: Are results believable taking study limitations into consideration?	Fair
Rebbeck, 2006[38]	Q1: Does the study population match the source population defined by the inclusion/exclusion criteria?"	Fair
	Q2: Was randomization adequately performed?	NR
	Q3: Was allocation adequately concealed?	Yes
	Q4: Did randomization result in equal practice and participant characteristics among study groups?	Yes
	Q5: Did the study maintain comparable groups?	Yes
	Q6: In cases of high loss to follow-up (or differential loss to follow-up), is the impact assessed (e.g., through sensitivity analysis or other adjustment method)?	NA
	Q7: Are outcomes assessed using valid and reliable measures?	Good
	Q8: Are outcomes implemented consistently across all study participants?	NR
	Q9: Were participants and intervention deliverers blinded during the study	Good
	Q10: Were the outcome assessors (e.g., data collectors) blinded to the intervention arm of participants?	NR
	Q11: Was the intervention delivered to all members of the intervention group as described?	NR
	Q12: Are the outcomes of interest pre-specified by the researchers?	Good
	Q13: Were the key confounding variables taken into account in the analysis and/or through randomization (e.g., successful randomization without loss to follow up, through matching, stratification, multivariate analysis, exclude based on confounding or other statistical adjustment)?	Yes
	Q14: Are the statistical methods used to assess the primary benefit outcomes appropriate to the data?	Yes
	Q15: Are results believable taking study limitations into consideration?	Good

Table C-2. Risk of bias for KQ 2 (continued)

Author, Year	Question	Response
Renzi, 2006[39]	Q1: Does the study population match the source population defined by the inclusion/exclusion criteria?"	Fair
	Q2: Was randomization adequately performed?	NR
	Q3: Was allocation adequately concealed?	NR
	Q4: Did randomization result in equal practice and participant characteristics among study groups?	Yes
	Q5: Did the study maintain comparable groups?	NR
	Q6: In cases of high loss to follow-up (or differential loss to follow-up), is the impact assessed (e.g., through sensitivity analysis or other adjustment method)?	NR
	Q7: Are outcomes assessed using valid and reliable measures?	NR
	Q8: Are outcomes implemented consistently across all study participants?	Good
	Q9: Were participants and intervention deliverers blinded during the study	NR
	Q10: Were the outcome assessors (e.g., data collectors) blinded to the intervention arm of participants?	NR
	Q11: Was the intervention delivered to all members of the intervention group as described?	NR
	Q12: Are the outcomes of interest pre-specified by the researchers?	Good
	Q13: Were the key confounding variables taken into account in the analysis and/or through randomization (e.g., successful randomization without loss to follow up, through matching, stratification, multivariate analysis, exclude based on confounding or other statistical adjustment)?	NR
	Q14: Are the statistical methods used to assess the primary benefit outcomes appropriate to the data?	Yes
	Q15: Are results believable taking study limitations into consideration?	Poor
Rimer, 2001[40]	Q1: Does the study population match the source population defined by the inclusion/exclusion criteria?"	Fair
	Q2: Was randomization adequately performed?	NR
	Q3: Was allocation adequately concealed?	NR
	Q4: Did randomization result in equal practice and participant characteristics among study groups?	Yes
	Q5: Did the study maintain comparable groups?	Yes
	Q6: In cases of high loss to follow-up (or differential loss to follow-up), is the impact assessed (e.g., through sensitivity analysis or other adjustment method)?	NA
	Q7: Are outcomes assessed using valid and reliable measures?	Fair
	Q8: Are outcomes implemented consistently across all study participants?	Good
	Q9: Were participants and intervention deliverers blinded during the study	NR
	Q10: Were the outcome assessors (e.g., data collectors) blinded to the intervention arm of participants?	NR
	Q11: Was the intervention delivered to all members of the intervention group as described?	Good
	Q12: Are the outcomes of interest pre-specified by the researchers?	Good
	Q13: Were the key confounding variables taken into account in the analysis and/or through randomization (e.g., successful randomization without loss to follow up, through matching, stratification, multivariate analysis, exclude based on confounding or other statistical adjustment)?	Yes
	Q14: Are the statistical methods used to assess the primary benefit outcomes appropriate to the data?	Yes
	Q15: Are results believable taking study limitations into consideration?	Fair

Table C-2. Risk of bias for KQ 2 (continued)

Author, Year	Question	Response
Simon, 2005[41]	Q1: Does the study population match the source population defined by the inclusion/exclusion criteria?"	Fair
	Q2: Was randomization adequately performed?	NR
	Q3: Was allocation adequately concealed?	No
	Q4: Did randomization result in equal practice and participant characteristics among study groups?	Yes
	Q5: Did the study maintain comparable groups?	NR
	Q6: In cases of high loss to follow-up (or differential loss to follow-up), is the impact assessed (e.g., through sensitivity analysis or other adjustment method)?	Yes
	Q7: Are outcomes assessed using valid and reliable measures?	Good
	Q8: Are outcomes implemented consistently across all study participants?	NR
	Q9: Were participants and intervention deliverers blinded during the study	Fair
	Q10: Were the outcome assessors (e.g., data collectors) blinded to the intervention arm of participants?	NA
	Q11: Was the intervention delivered to all members of the intervention group as described?	Fair
	Q12: Are the outcomes of interest pre-specified by the researchers?	Good
	Q13: Were the key confounding variables taken into account in the analysis and/or through randomization (e.g., successful randomization without loss to follow up, through matching, stratification, multivariate analysis, exclude based on confounding or other statistical adjustment)?	Yes
	Q14: Are the statistical methods used to assess the primary benefit outcomes appropriate to the data?	Yes
	Q15: Are results believable taking study limitations into consideration?	Good
Soler, 2010[42]	Q1: Does the study population match the source population defined by the inclusion/exclusion criteria?"	NR
	Q2: Was randomization adequately performed?	NR
	Q3: Was allocation adequately concealed?	NR
	Q4: Did randomization result in equal practice and participant characteristics among study groups?	Yes
	Q5: Did the study maintain comparable groups?	Yes
	Q6: In cases of high loss to follow-up (or differential loss to follow-up), is the impact assessed (e.g., through sensitivity analysis or other adjustment method)?	NR
	Q7: Are outcomes assessed using valid and reliable measures?	Good
	Q8: Are outcomes implemented consistently across all study participants?	Fair
	Q9: Were participants and intervention deliverers blinded during the study	NR
	Q10: Were the outcome assessors (e.g., data collectors) blinded to the intervention arm of participants?	NR
	Q11: Was the intervention delivered to all members of the intervention group as described?	NR
	Q12: Are the outcomes of interest pre-specified by the researchers?	Fair
	Q13: Were the key confounding variables taken into account in the analysis and/or through randomization (e.g., successful randomization without loss to follow up, through matching, stratification, multivariate analysis, exclude based on confounding or other statistical adjustment)?	Yes
	Q14: Are the statistical methods used to assess the primary benefit outcomes appropriate to the data?	Yes
	Q15: Are results believable taking study limitations into consideration?	Fair

Table C-2. Risk of bias for KQ 2 (continued)

Author, Year	Question	Response
Sullivan, 2010[43]	Q1: Does the study population match the source population defined by the inclusion/exclusion criteria?"	NR
	Q2: Was randomization adequately performed?	NR
	Q3: Was allocation adequately concealed?	NR
	Q4: Did randomization result in equal practice and participant characteristics among study groups?	Yes
	Q5: Did the study maintain comparable groups?	No
	Q6: In cases of high loss to follow-up (or differential loss to follow-up), is the impact assessed (e.g., through sensitivity analysis or other adjustment method)?	Yes
	Q7: Are outcomes assessed using valid and reliable measures?	Good
	Q8: Are outcomes implemented consistently across all study participants?	Good
	Q9: Were participants and intervention deliverers blinded during the study	NR
	Q10: Were the outcome assessors (e.g., data collectors) blinded to the intervention arm of participants?	NR
	Q11: Was the intervention delivered to all members of the intervention group as described?	NR
	Q12: Are the outcomes of interest pre-specified by the researchers?	Good
	Q13: Were the key confounding variables taken into account in the analysis and/or through randomization (e.g., successful randomization without loss to follow up, through matching, stratification, multivariate analysis, exclude based on confounding or other statistical adjustment)?	Yes
	Q14: Are the statistical methods used to assess the primary benefit outcomes appropriate to the data?	Yes
	Q15: Are results believable taking study limitations into consideration?	Fair
Van Driel, 2007[44]	Q1: Does the study population match the source population defined by the inclusion/exclusion criteria?"	Fair
	Q2: Was randomization adequately performed?	NR
	Q3: Was allocation adequately concealed?	NR
	Q4: Did randomization result in equal practice and participant characteristics among study groups?	Yes
	Q5: Did the study maintain comparable groups?	No
	Q6: In cases of high loss to follow-up (or differential loss to follow-up), is the impact assessed (e.g., through sensitivity analysis or other adjustment method)?	Yes
	Q7: Are outcomes assessed using valid and reliable measures?	NR
	Q8: Are outcomes implemented consistently across all study participants?	NR
	Q9: Were participants and intervention deliverers blinded during the study	NR
	Q10: Were the outcome assessors (e.g., data collectors) blinded to the intervention arm of participants?	Yes
	Q11: Was the intervention delivered to all members of the intervention group as described?	NR
	Q12: Are the outcomes of interest pre-specified by the researchers?	Good
	Q13: Were the key confounding variables taken into account in the analysis and/or through randomization (e.g., successful randomization without loss to follow up, through matching, stratification, multivariate analysis, exclude based on confounding or other statistical adjustment)?	Yes
	Q14: Are the statistical methods used to assess the primary benefit outcomes appropriate to the data?	Yes
	Q15: Are results believable taking study limitations into consideration?	Poor

Table C-2. Risk of bias for KQ 2 (continued)

Author, Year	Question	Response
van Steenkiste, 2007[45]	Q1: Does the study population match the source population defined by the inclusion/exclusion criteria?"	Fair
	Q2: Was randomization adequately performed?	Yes
	Q3: Was allocation adequately concealed?	NR
	Q4: Did randomization result in equal practice and participant characteristics among study groups?	No
	Q5: Did the study maintain comparable groups?	No
	Q6: In cases of high loss to follow-up (or differential loss to follow-up), is the impact assessed (e.g., through sensitivity analysis or other adjustment method)?	NR
	Q7: Are outcomes assessed using valid and reliable measures?	NR
	Q8: Are outcomes implemented consistently across all study participants?	NR
	Q9: Were participants and intervention deliverers blinded during the study	NR
	Q10: Were the outcome assessors (e.g., data collectors) blinded to the intervention arm of participants?	NR
	Q11: Was the intervention delivered to all members of the intervention group as described?	Poor
	Q12: Are the outcomes of interest pre-specified by the researchers?	Good
	Q13: Were the key confounding variables taken into account in the analysis and/or through randomization (e.g., successful randomization without loss to follow up, through matching, stratification, multivariate analysis, exclude based on confounding or other statistical adjustment)?	Yes
	Q14: Are the statistical methods used to assess the primary benefit outcomes appropriate to the data?	Yes
	Q15: Are results believable taking study limitations into consideration?	Poor
Varonen, 2007[46]	Q1: Does the study population match the source population defined by the inclusion/exclusion criteria?"	NR
	Q2: Was randomization adequately performed?	No
	Q3: Was allocation adequately concealed?	NR
	Q4: Did randomization result in equal practice and participant characteristics among study groups?	NR
	Q5: Did the study maintain comparable groups?	No
	Q6: In cases of high loss to follow-up (or differential loss to follow-up), is the impact assessed (e.g., through sensitivity analysis or other adjustment method)?	NR
	Q7: Are outcomes assessed using valid and reliable measures?	NR
	Q8: Are outcomes implemented consistently across all study participants?	NR
	Q9: Were participants and intervention deliverers blinded during the study	NR
	Q10: Were the outcome assessors (e.g., data collectors) blinded to the intervention arm of participants?	NR
	Q11: Was the intervention delivered to all members of the intervention group as described?	NR
	Q12: Are the outcomes of interest pre-specified by the researchers?	Good
	Q13: Were the key confounding variables taken into account in the analysis and/or through randomization (e.g., successful randomization without loss to follow up, through matching, stratification, multivariate analysis, exclude based on confounding or other statistical adjustment)?	NR
	Q14: Are the statistical methods used to assess the primary benefit outcomes appropriate to the data?	Yes
	Q15: Are results believable taking study limitations into consideration?	Poor

Table C-2. Risk of bias for KQ 2

Author, Year	Question	Response
Wadland, 2001[47]	Q1: Does the study population match the source population defined by the inclusion/exclusion criteria?"	NR
	Q2: Was randomization adequately performed?	NR
	Q3: Was allocation adequately concealed?	NR
	Q4: Did randomization result in equal practice and participant characteristics among study groups?	Yes
	Q5: Did the study maintain comparable groups?	No
	Q6: In cases of high loss to follow-up (or differential loss to follow-up), is the impact assessed (e.g., through sensitivity analysis or other adjustment method)?	Yes
	Q7: Are outcomes assessed using valid and reliable measures?	Good
	Q8: Are outcomes implemented consistently across all study participants?	Poor
	Q9: Were participants and intervention deliverers blinded during the study	NR
	Q10: Were the outcome assessors (e.g., data collectors) blinded to the intervention arm of participants?	NR
	Q11: Was the intervention delivered to all members of the intervention group as described?	NR
	Q12: Are the outcomes of interest pre-specified by the researchers?	Good
	Q13: Were the key confounding variables taken into account in the analysis and/or through randomization (e.g., successful randomization without loss to follow up, through matching, stratification, multivariate analysis, exclude based on confounding or other statistical adjustment)?	Yes
	Q14: Are the statistical methods used to assess the primary benefit outcomes appropriate to the data?	Yes
	Q15: Are results believable taking study limitations into consideration?	Poor
Watson, 2002[48]	Q1: Does the study population match the source population defined by the inclusion/exclusion criteria?"	Fair
	Q2: Was randomization adequately performed?	Yes
	Q3: Was allocation adequately concealed?	NR
	Q4: Did randomization result in equal practice and participant characteristics among study groups?	Yes
	Q5: Did the study maintain comparable groups?	Yes
	Q6: In cases of high loss to follow-up (or differential loss to follow-up), is the impact assessed (e.g., through sensitivity analysis or other adjustment method)?	NA
	Q7: Are outcomes assessed using valid and reliable measures?	Poor
	Q8: Are outcomes implemented consistently across all study participants?	Fair
	Q9: Were participants and intervention deliverers blinded during the study	NR
	Q10: Were the outcome assessors (e.g., data collectors) blinded to the intervention arm of participants?	No
	Q11: Was the intervention delivered to all members of the intervention group as described?	Good
	Q12: Are the outcomes of interest pre-specified by the researchers?	Good
	Q13: Were the key confounding variables taken into account in the analysis and/or through randomization (e.g., successful randomization without loss to follow up, through matching, stratification, multivariate analysis, exclude based on confounding or other statistical adjustment)?	Yes
	Q14: Are the statistical methods used to assess the primary benefit outcomes appropriate to the data?	Yes
	Q15: Are results believable taking study limitations into consideration?	Fair

Table C-2. Risk of bias for KQ 2 (continued)

Author, Year	Question	Response
Wetter, 2006[49]	Q1: Does the study population match the source population defined by the inclusion/exclusion criteria?"	Fair
	Q2: Was randomization adequately performed?	NR
	Q3: Was allocation adequately concealed?	NR
	Q4: Did randomization result in equal practice and participant characteristics among study groups?	Yes
	Q5: Did the study maintain comparable groups?	Yes
	Q6: In cases of high loss to follow-up (or differential loss to follow-up), is the impact assessed (e.g., through sensitivity analysis or other adjustment method)?	NA
	Q7: Are outcomes assessed using valid and reliable measures?	NR
	Q8: Are outcomes implemented consistently across all study participants?	Good
	Q9: Were participants and intervention deliverers blinded during the study	NR
	Q10: Were the outcome assessors (e.g., data collectors) blinded to the intervention arm of participants?	NA
	Q11: Was the intervention delivered to all members of the intervention group as described?	Fair
	Q12: Are the outcomes of interest pre-specified by the researchers?	Good
	Q13: Were the key confounding variables taken into account in the analysis and/or through randomization (e.g., successful randomization without loss to follow up, through matching, stratification, multivariate analysis, exclude based on confounding or other statistical adjustment)?	Yes
	Q14: Are the statistical methods used to assess the primary benefit outcomes appropriate to the data?	Yes
	Q15: Are results believable taking study limitations into consideration?	Good
Wolters, 2005[50]	Q1: Does the study population match the source population defined by the inclusion/exclusion criteria?"	Yes
	Q2: Was randomization adequately performed?	Yes
	Q3: Was allocation adequately concealed?	Yes
	Q4: Did randomization result in equal practice and participant characteristics among study groups?	Yes
	Q5: Did the study maintain comparable groups?	No
	Q6: In cases of high loss to follow-up (or differential loss to follow-up), is the impact assessed (e.g., through sensitivity analysis or other adjustment method)?	Yes
	Q7: Are outcomes assessed using valid and reliable measures?	Good
	Q8: Are outcomes implemented consistently across all study participants?	NR
	Q9: Were participants and intervention deliverers blinded during the study	Good
	Q10: Were the outcome assessors (e.g., data collectors) blinded to the intervention arm of participants?	NR
	Q11: Was the intervention delivered to all members of the intervention group as described?	NR
	Q12: Are the outcomes of interest pre-specified by the researchers?	Good
	Q13: Were the key confounding variables taken into account in the analysis and/or through randomization (e.g., successful randomization without loss to follow up, through matching, stratification, multivariate analysis, exclude based on confounding or other statistical adjustment)?	Yes
	Q14: Are the statistical methods used to assess the primary benefit outcomes appropriate to the data?	Yes
	Q15: Are results believable taking study limitations into consideration?	Good

Table C-2. Risk of bias for KQ 2 (continued)

Author, Year	Question	Response
Wright, 2008[51]	Q1: Does the study population match the source population defined by the inclusion/exclusion criteria?"	Fair
	Q2: Was randomization adequately performed?	Yes
	Q3: Was allocation adequately concealed?	NR
	Q4: Did randomization result in equal practice and participant characteristics among study groups?	Yes
	Q5: Did the study maintain comparable groups?	Yes
	Q6: In cases of high loss to follow-up (or differential loss to follow-up), is the impact assessed (e.g., through sensitivity analysis or other adjustment method)?	NA
	Q7: Are outcomes assessed using valid and reliable measures?	Good
	Q8: Are outcomes implemented consistently across all study participants?	Good
	Q9: Were participants and intervention deliverers blinded during the study	Fair
	Q10: Were the outcome assessors (e.g., data collectors) blinded to the intervention arm of participants?	NA
	Q11: Was the intervention delivered to all members of the intervention group as described?	Good
	Q12: Are the outcomes of interest pre-specified by the researchers?	Good
	Q13: Were the key confounding variables taken into account in the analysis and/or through randomization (e.g., successful randomization without loss to follow up, through matching, stratification, multivariate analysis, exclude based on confounding or other statistical adjustment)?	Yes
	Q14: Are the statistical methods used to assess the primary benefit outcomes appropriate to the data?	Yes
	Q15: Are results believable taking study limitations into consideration?	Good

Table C-3. Risk of bias for KQ 3

Author, Year	Question	Response
Akl, 2012[52]	Q1: Does the study population match the source population defined by the inclusion/exclusion criteria?"	Fair
	Q2: Was randomization adequately performed?	NR
	Q3: Was allocation adequately concealed?	NR
	Q4: Did randomization result in equal practice and participant characteristics among study groups?	Yes
	Q5: Did the study maintain comparable groups?	Yes
	Q6: In cases of high loss to follow-up (or differential loss to follow-up), is the impact assessed (e.g., through sensitivity analysis or other adjustment method)?	NA
	Q7: Are outcomes assessed using valid and reliable measures?	Fair
	Q8: Are outcomes implemented consistently across all study participants?	Good
	Q9: Were participants and intervention deliverers blinded during the study	NR
	Q10: Were the outcome assessors (e.g., data collectors) blinded to the intervention arm of participants?	NR
	Q11: Was the intervention delivered to all members of the intervention group as described?	Good
	Q12: Are the outcomes of interest pre-specified by the researchers?	Good
	Q13: Were the key confounding variables taken into account in the analysis and/or through randomization (e.g., successful randomization without loss to follow up, through matching, stratification, multivariate analysis, exclude based on confounding or other statistical adjustment)?	Yes
	Q14: Are the statistical methods used to assess the primary benefit outcomes appropriate to the data?	Yes
	Q15: Are results believable taking study limitations into consideration?	Fair
Brewer, 2011[53]	Q1: Does the study population match the source population defined by the inclusion/exclusion criteria?"	Poor
	Q2: Was randomization adequately performed?	NA
	Q3: Was allocation adequately concealed?	NR
	Q4: Did randomization result in equal practice and participant characteristics among study groups?	NA
	Q5: Did the study maintain comparable groups?	Yes
	Q6: In cases of high loss to follow-up (or differential loss to follow-up), is the impact assessed (e.g., through sensitivity analysis or other adjustment method)?	NA
	Q7: Are outcomes assessed using valid and reliable measures?	Fair
	Q8: Are outcomes implemented consistently across all study participants?	Good
	Q9: Were participants and intervention deliverers blinded during the study	NR
	Q10: Were the outcome assessors (e.g., data collectors) blinded to the intervention arm of participants?	NR
	Q11: Was the intervention delivered to all members of the intervention group as described?	Good
	Q12: Are the outcomes of interest pre-specified by the researchers?	Good
	Q13: Were the key confounding variables taken into account in the analysis and/or through randomization (e.g., successful randomization without loss to follow up, through matching, stratification, multivariate analysis, exclude based on confounding or other statistical adjustment)?	Yes
	Q14: Are the statistical methods used to assess the primary benefit outcomes appropriate to the data?	Yes
	Q15: Are results believable taking study limitations into consideration?	Fair

Table C-3. Risk of bias for KQ 3 (continued)

Author, Year	Question	Response
Han, 2011[54]	Q1: Does the study population match the source population defined by the inclusion/exclusion criteria?"	Fair
	Q2: Was randomization adequately performed?	NR
	Q3: Was allocation adequately concealed?	NR
	Q4: Did randomization result in equal practice and participant characteristics among study groups?	Yes
	Q5: Did the study maintain comparable groups?	Yes
	Q6: In cases of high loss to follow-up (or differential loss to follow-up), is the impact assessed (e.g., through sensitivity analysis or other adjustment method)?	NA
	Q7: Are outcomes assessed using valid and reliable measures?	Good
	Q8: Are outcomes implemented consistently across all study participants?	Good
	Q9: Were participants and intervention deliverers blinded during the study	NR
	Q10: Were the outcome assessors (e.g., data collectors) blinded to the intervention arm of participants?	NR
	Q11: Was the intervention delivered to all members of the intervention group as described?	Good
	Q12: Are the outcomes of interest pre-specified by the researchers?	Good
	Q13: Were the key confounding variables taken into account in the analysis and/or through randomization (e.g., successful randomization without loss to follow up, through matching, stratification, multivariate analysis, exclude based on confounding or other statistical adjustment)?	Yes
	Q14: Are the statistical methods used to assess the primary benefit outcomes appropriate to the data?	Yes
	Q15: Are results believable taking study limitations into consideration?	Fair
Johnson, 1995[55]	Q1: Does the study population match the source population defined by the inclusion/exclusion criteria?"	Poor
	Q2: Was randomization adequately performed?	NR
	Q3: Was allocation adequately concealed?	NR
	Q4: Did randomization result in equal practice and participant characteristics among study groups?	NR
	Q5: Did the study maintain comparable groups?	NR
	Q6: In cases of high loss to follow-up (or differential loss to follow-up), is the impact assessed (e.g., through sensitivity analysis or other adjustment method)?	NA
	Q7: Are outcomes assessed using valid and reliable measures?	Fair
	Q8: Are outcomes implemented consistently across all study participants?	Fair
	Q9: Were participants and intervention deliverers blinded during the study	NR
	Q10: Were the outcome assessors (e.g., data collectors) blinded to the intervention arm of participants?	NR
	Q11: Was the intervention delivered to all members of the intervention group as described?	Good
	Q12: Are the outcomes of interest pre-specified by the researchers?	Good
	Q13: Were the key confounding variables taken into account in the analysis and/or through randomization (e.g., successful randomization without loss to follow up, through matching, stratification, multivariate analysis, exclude based on confounding or other statistical adjustment)?	No
	Q14: Are the statistical methods used to assess the primary benefit outcomes appropriate to the data?	No
	Q15: Are results believable taking study limitations into consideration?	Poor

Table C-3. Risk of bias for KQ 3 (continued)

Author, Year	Question	Response
McCormack, 2011[56]	Q1: Does the study population match the source population defined by the inclusion/exclusion criteria?"	Yes
	Q2: Was randomization adequately performed?	No
	Q3: Was allocation adequately concealed?	No
	Q4: Did randomization result in equal practice and participant characteristics among study groups?	No
	Q5: Did the study maintain comparable groups?	NR
	Q6: In cases of high loss to follow-up (or differential loss to follow-up), is the impact assessed (e.g., through sensitivity analysis or other adjustment method)?	Yes
	Q7: Are outcomes assessed using valid and reliable measures?	Good
	Q8: Are outcomes implemented consistently across all study participants?	Good
	Q9: Were participants and intervention deliverers blinded during the study	Fair
	Q10: Were the outcome assessors (e.g., data collectors) blinded to the intervention arm of participants?	NR
	Q11: Was the intervention delivered to all members of the intervention group as described?	Fair
	Q12: Are the outcomes of interest pre-specified by the researchers?	Good
	Q13: Were the key confounding variables taken into account in the analysis and/or through randomization (e.g., successful randomization without loss to follow up, through matching, stratification, multivariate analysis, exclude based on confounding or other statistical adjustment)?	Yes
	Q14: Are the statistical methods used to assess the primary benefit outcomes appropriate to the data?	Yes
	Q15: Are results believable taking study limitations into consideration?	Fair
Perneger, 2010[57]	Q1: Does the study population match the source population defined by the inclusion/exclusion criteria?"	Fair
	Q2: Was randomization adequately performed?	NR
	Q3: Was allocation adequately concealed?	NR
	Q4: Did randomization result in equal practice and participant characteristics among study groups?	NR
	Q5: Did the study maintain comparable groups?	NR
	Q6: In cases of high loss to follow-up (or differential loss to follow-up), is the impact assessed (e.g., through sensitivity analysis or other adjustment method)?	NA
	Q7: Are outcomes assessed using valid and reliable measures?	Fair
	Q8: Are outcomes implemented consistently across all study participants?	NR
	Q9: Were participants and intervention deliverers blinded during the study	NR
	Q10: Were the outcome assessors (e.g., data collectors) blinded to the intervention arm of participants?	NR
	Q11: Was the intervention delivered to all members of the intervention group as described?	Good
	Q12: Are the outcomes of interest pre-specified by the researchers?	Good
	Q13: Were the key confounding variables taken into account in the analysis and/or through randomization (e.g., successful randomization without loss to follow up, through matching, stratification, multivariate analysis, exclude based on confounding or other statistical adjustment)?	Yes
	Q14: Are the statistical methods used to assess the primary benefit outcomes appropriate to the data?	Yes
	Q15: Are results believable taking study limitations into consideration?	Fair

Table C-3. Risk of bias for KQ 3 (continued)

Author, Year	Question	Response
Schwartz, 2011[58]	Q1: Does the study population match the source population defined by the inclusion/exclusion criteria?"	Fair
	Q2: Was randomization adequately performed?	Yes
	Q3: Was allocation adequately concealed?	Yes
	Q4: Did randomization result in equal practice and participant characteristics among study groups?	Yes
	Q5: Did the study maintain comparable groups?	Yes
	Q6: In cases of high loss to follow-up (or differential loss to follow-up), is the impact assessed (e.g., through sensitivity analysis or other adjustment method)?	NA
	Q7: Are outcomes assessed using valid and reliable measures?	Fair
	Q8: Are outcomes implemented consistently across all study participants?	Good
	Q9: Were participants and intervention deliverers blinded during the study	Fair
	Q10: Were the outcome assessors (e.g., data collectors) blinded to the intervention arm of participants?	Yes
	Q11: Was the intervention delivered to all members of the intervention group as described?	Good
	Q12: Are the outcomes of interest pre-specified by the researchers?	Good
	Q13: Were the key confounding variables taken into account in the analysis and/or through randomization (e.g., successful randomization without loss to follow up, through matching, stratification, multivariate analysis, exclude based on confounding or other statistical adjustment)?	Yes
	Q14: Are the statistical methods used to assess the primary benefit outcomes appropriate to the data?	Yes
	Q15: Are results believable taking study limitations into consideration?	Good

References

1. Cox D, Cox AD. Communicating the consequences of early detection: The role of evidence and framing. J Market. 2001;65(3):91-103.
2. Elder JP, Ayala GX, Campbell NR, et al. Long-term effects of a communication intervention for Spanish-dominant Latinas. Am J Prev Med. 2006 Aug;31(2):159-66. PMID: 16829333.
3. Jibaja-Weiss ML, Volk RJ, Kingery P, et al. Tailored messages for breast and cervical cancer screening of low-income and minority women using medical records data. Patient Educ Couns. 2003 Jun;50(2):123-32. PMID: 12781927.
4. Schneider TR, Salovey P, Apanovitch AM, et al. The effects of message framing and ethnic targeting on mammography use among low-income women. Health Psychol. 2001;20(4):256-66. PMID: 11515737.
5. Vernon SW, del Junco DJ, Tiro JA, et al. Promoting regular mammography screening II. Results from a randomized controlled trial in US women veterans. J Natl Cancer Inst. 2008 Mar 5;100(5):347-58. PMID: 18314474.
6. Bahrami M, Deery C, Clarkson JE, et al. Effectiveness of strategies to disseminate and implement clinical guidelines for the management of impacted and unerupted third molars in primary dental care, a cluster randomised controlled trial. Br Dent J. 2004 Dec 11;197(11):691-6; discussion 88. PMID: 15592551.
7. Banait G, Sibbald B, Thompson D, et al. Modifying dyspepsia management in primary care: a cluster randomised controlled trial of educational outreach compared with passive guideline dissemination. Br J Gen Pract. 2003 Feb;53(487):94-100. PMID: 12817353.
8. Beaulieu MD, Brophy J, Jacques A, et al. Drug treatment of stable angina pectoris and mass dissemination of therapeutic guidelines: a randomized controlled trial. QJM. 2004 Jan;97(1):21-31. PMID: 14702508.
9. Becker A, Leonhardt C, Kochen MM, et al. Effects of two guideline implementation strategies on patient outcomes in primary care: a cluster randomized controlled trial. Spine (Phila Pa 1976). 2008(5):473-80. PMID: 18317189.
10. Bekkering GE, Hendriks HJ, van Tulder MW, et al. Effect on the process of care of an active strategy to implement clinical guidelines on physiotherapy for low back pain: a cluster randomised controlled trial. Qual Saf Health Care. 2005(2):107-12. PMID: 15805455.
11. Bekkering GE, van Tulder MW, Hendriks EJ, et al. Implementation of clinical guidelines on physical therapy for patients with low back pain: randomized trial comparing patient outcomes after a standard and active implementation strategy. Phys Ther. 2005(6):544-55. PMID: 15921475.
12. Bishop PB, Wing PC. Knowledge transfer in family physicians managing patients with acute low back pain: a prospective randomized control trial. Spine J. 2006 May-Jun;6(3):282-8. PMID: 16651222.
13. Campbell MK, James A, Hudson MA, et al. Improving multiple behaviors for colorectal cancer prevention among African American church members. Health Psychol. 2004;23(5):492-502. PMID: 15367069.
14. Carney PA, Harwood BG, Greene MA, et al. Impact of a telephone counseling intervention on transitions in stage of change and adherence to interval mammography screening (United States). Cancer Causes Control. 2005 Sep;16(7):799-807. PMID: 16132790.
15. Christakis DA, Zimmerman FJ, Rivara FP, et al. Improving pediatric prevention via the internet: a randomized, controlled trial. Pediatrics. 2006 Sep;118(3):1157-66. PMID: 16951011.
16. Church TR, Yeazel MW, Jones RM, et al. A randomized trial of direct mailing of fecal occult blood tests to increase colorectal cancer screening. J Natl Cancer Inst. 2004 May 19;96(10):770-80. PMID: 15150305.

17. Daniels EC, Bacon J, Denisio S, et al. Translation squared: improving asthma care for high-disparity populations through a safety net practice-based research network. J Asthma. 2005 Jul-Aug;42(6):499-505. PMID: 16293546.

18. Davis J, Roberts R, Davidson DL, et al. Implementation strategies for a Scottish national epilepsy guideline in primary care: results of the Tayside Implementation of Guidelines in Epilepsy Randomized (TIGER) trial. Epilepsia. 2004 Jan;45(1):28-34. PMID: 14692904.

19. Eaton CB, Parker DR, Borkan J, et al. Translating cholesterol guidelines into primary care practice: a multimodal cluster randomized trial. Ann Fam Med. 2011 Nov-Dec;9(6):528-37. PMID: 22084264.

20. Feldstein A, Elmer PJ, Smith DH, et al. Electronic medical record reminder improves osteoporosis management after a fracture: a randomized, controlled trial. J Am Geriatr Soc. 2006 Mar;54(3):450-7. PMID: 16551312.

21. Figueiras A, Sastre I, Tato F, et al. One-to-one versus group sessions to improve prescription in primary care: a pragmatic randomized controlled trial. Med Care. 2001 Feb;39(2):158-67. PMID: 11176553.

22. Gattellari M, Ward JE. A community-based randomised controlled trial of three different educational resources for men about prostate cancer screening. Patient Educ Couns. 2005;57(2):162-82. PMID: 15911190.

23. Hagmolen of ten Have W, van den Berg NJ, van der Palen J, et al. Implementation of an asthma guideline for the management of childhood asthma in general practice: a randomised controlled trial. Prim Care Respir J. 2008 Jun;17(2):90-6. PMID: 18274692.

24. Jain MK, Heyland D, Dhaliwal R, et al. Dissemination of the Canadian clinical practice guidelines for nutrition support: results of a cluster randomized controlled trial. Crit Care Med. 2006 Sep;34(9):2362-9. PMID: 16850001.

25. Jousimaa J, Mäkelä M, Kunnamo I, et al. Primary care guidelines on consultation practices: the effectiveness of computerized versus paper-based versions. A cluster randomized controlled trial among newly qualified primary care physicians. Int J Technol Assess Health Care. 2002 Sum(3):586-96. PMID: 12391951.

26. Junghans C, Feder G, Timmis AD, et al. Effect of patient-specific ratings vs conventional guidelines on investigation decisions in angina: Appropriateness of Referral and Investigation in Angina (ARIA) Trial. Arch Intern Med. 2007 Jan 22;167(2):195-202. PMID: 17242322.

27. Kennedy ADM, Sculpher MJ, Coulter A, et al. A multicentre randomised controlled trial assessing the costs and benefits of using structured information and analysis of women's preferences in the management of menorrhagia. Health Technol Assess. 2003;7(8):1-76. PMID: 12633525.

28. King AC, Friedman R, Marcus B, et al. Ongoing physical activity advice by humans versus computers: The Community Health Advice by Telephone (CHAT) trial. Health Psychol. 2007 Nov;26(6):718-27. PMID: 18020844.

29. Lenza M, Belloti Joao C, Gomes dos Santos Joao B, et al. Surgical interventions for treating acute fractures or non-union of the middle third of the clavicle. Cochrane Database of Systematic Reviews: John Wiley & Sons, Ltd; 2009.

30. Lien LF, Brown AJ, Ard JD, et al. Effects of PREMIER lifestyle modifications on participants with and without the metabolic syndrome. Hypertension. 2007 Oct;50(4):609-16. PMID: 17698724.

31. Manfredi C, Cho YI, Warnecke R, et al. Dissemination strategies to improve implementation of the PHS smoking cessation guideline in MCH public health clinics: experimental evaluation results and contextual factors. Health Educ Res. 2011 Apr;26(2):348-60. PMID: 21398375.

32. Marcus BH, Napolitano MA, King AC, et al. Telephone versus print delivery of an individualized motivationally tailored physical activity intervention: Project STRIDE. Health Psychol. 2007 Jul;26(4):401-9. PMID: 17605559.

33. Maxwell AE, Bastani R, Danao LL, et al. Results of a community-based randomized trial to increase colorectal cancer screening among Filipino Americans. Am J Public Health. 2010 Nov;100(11):2228-34. PMID: 20864724.

34. Murtaugh CM, Pezzin LE, McDonald MV, et al. Just-in-time evidence-based e-mail "reminders" in home health care: impact on nurse practices. Health Serv Res. 2005 Jun;40(3):849-64. PMID: 15960694.

35. Paradis HA, Conn KM, Gewirtz JR, et al. Innovative delivery of newborn anticipatory guidance: a randomized, controlled trial incorporating media-based learning into primary care. Acad Pediatr. 2011 Jan-Feb;11(1):27-33. PMID: 21272821.

36. Partin MR, Nelson D, Radosevich D, et al. Randomized trial examining the effect of two prostate cancer screening educational interventions on patient knowledge, preferences, and behaviors. J Gen Intern Med. 2004 Aug;19(8):835-42. PMID: 15242468.

37. Rahme E, Choquette D, Beaulieu M, et al. Impact of a general practitioner educational intervention on osteoarthritis treatment in an elderly population. Am J Med. 2005 Nov;118(11):1262-70. PMID: 16271911.

38. Rebbeck T, Maher CG, Refshauge KM. Evaluating two implementation strategies for whiplash guidelines in physiotherapy: a cluster randomised trial. Aust J Physiother. 2006;52(3):165-74. PMID: 16942451.

39. Renzi PM, Ghezzo H, Goulet S, et al. Paper stamp checklist tool enhances asthma guidelines knowledge and implementation by primary care physicians. Can Respir J. 2006 May-Jun;13(4):193-7. PMID: 16779463.

40. Han PK, Klein WM, Arora NK. Varieties of uncertainty in health care: a conceptual taxonomy. Med Decis Making. 2011 Nov-Dec;31(6):828-38. PMID: 22067431.

41. Simon SR, Majumdar SR, Prosser LA, et al. Group versus individual academic detailing to improve the use of antihypertensive medications in primary care: a cluster-randomized controlled trial. Am J Med. 2005 May;118(5):521-8. PMID: 15866255.

42. Soler N, Ballester E, Martin A, et al. Changes in management of chronic obstructive pulmonary disease (COPD) in primary care: EMMEPOC study. Respir Med. 2010 Jan;104(1):67-75. PMID: 20122630.

43. Sullivan MD, Gaster B, Russo J, et al. Randomized trial of web-based training about opioid therapy for chronic pain. Clin J Pain. 2010 Jul;26(6):512-7. PMID: 20551726.

44. van Driel ML, Coenen S, Dirven K, et al. What is the role of quality circles in strategies to optimise antibiotic prescribing? A pragmatic cluster-randomised controlled trial in primary care. Qual Saf Health Care. 2007 Jun;16(3):197-202. PMID: 17545346.

45. van Steenkiste B, van der Weijden T, Stoffers HE, et al. Improving cardiovascular risk management: a randomized, controlled trial on the effect of a decision support tool for patients and physicians. Eur J Cardiovasc Prev Rehabil. 2007 Feb;14(1):44-50. PMID: 17301626.

46. Varonen H, Rautakorpi UM, Nyberg S, et al. Implementing guidelines on acute maxillary sinusitis in general practice--a randomized controlled trial. Fam Pract. 2007 Apr;24(2):201-6. PMID: 17237494.

47. Wadland WC, Soffelmayr B, Ives K. Enhancing smoking cessation of low-income smokers in managed care. J Fam Pract. 2001 Feb;50(2):138-44. PMID: 11219562.

48. Watson MC, Bond CM, Grimshaw JM, et al. Educational strategies to promote evidence-based community pharmacy practice: a cluster randomized controlled trial (RCT). Fam Pract. 2002 Oct;19(5):529-36. PMID: 12356708.

49. Wetter DW, Mazas C, Daza P, et al. Reaching and treating Spanish-speaking smokers through the National Cancer Institute's Cancer Information Service. A randomized controlled trial. Cancer. 2007 Jan 15;109(2 Suppl):406-13. PMID: 17149758.

50. Wolters R, Wensing M, Klomp M, et al. Effects of distance learning on clinical management of LUTS in primary care: a randomised trial. Patient Educ Couns. 2005 Nov;59(2):212-8. PMID: 16257627.

51. Wright FC, Gagliardi AR, Law CH, et al. A randomized controlled trial to improve lymph node assessment in stage II colon cancer. Arch Surg. 2008 Nov;143(11):1050-5; discussion 5. PMID: 19015462.

52. Akl EA, Guyatt GH, Irani J, et al. "Might" or "suggest"? No wording approach was clearly superior in conveying the strength of recommendation. J Clin Epidemiol. 2012 Mar;65(3):268-75. PMID: 22075112.

53. Brewer NT, Richman AR, DeFrank JT, et al. Improving communication of breast cancer recurrence risk. Breast Cancer Res Treat. 2012 Jun;133(2):553-61. PMID: 21964579.

54. Han PKJ, Klein WMP, Lehman T, et al. Communication of uncertainty regarding individualized cancer risk estimates: effects and influential factors. Med Decis Making. 2011 Mar-Apr;31(2):354-66. PMID: 20671211.

55. Johnson BB, Slovic P. Presenting uncertainty in health risk assessment: Initial studies of its effects on risk perception and trust. Risk Anal. 1995 Aug;15(4):485-94. PMID: 7480948.

56. McCormack L, Treiman K, Bann C, et al. Translating medical evidence to promote informed health care decisions. Health Serv Res. 2011 Aug;46(4):1200-23. PMID: 21352225.

57. Perneger TV, Schiesari L, Cullati S, et al. Does information about risks and benefits improve the decision-making process in cancer screening - randomized study. Cancer Epidemiol. 2011 Dec;35(6):574-9. PMID: 21622043.

58. Schwartz LM, Woloshin S. Communicating uncertainties about prescription drugs to the public: a national randomized trial. Arch Intern Med. 2011 Sep 12;171(16):1463-8. PMID: 21911629.

Appendix D. Evidence Tables for Key Question 1

Table D-1. Key question 1 study design details

Author, Year	Research Objective	Funding Source	Geographic Location, Setting Type, Setting Description	Study Design	Primary Outcomes	Measurement Intervals	Other Notes
Cox 2001[1]	Examine how consumers' beliefs and attitudes toward screening are affected by two specific message-design factors: (1) whether screening consequences are communicated with anecdotal evidence or statistical evidence and (2) whether these consequences are framed in terms of potential losses or potential gains.	Academic	US Other Social and volunteer organizations in a Midwestern metropolitan area.	RCT	Perceived likelihood of getting a memmogram	immediate post-test	
Elder 2005[2]	The present study examined two innovative lifestyle behavior change approaches to reduce dietary fat and to increase fiber. Analyses emphasized (a) whether personalized counseling via promotora plus tailored print materials used in an interactive format were more effective than tailored materials delivered in a distance learning format, and (b) whether these two innovations were more effective than standard off-the-shelf materials TARGETED (culturally) to a Latino population (controls).	Government	US Community-based settings San Diego County, with dominant Latino populations	RCT	Percent calories from fat Number of daily grams of fiber	Baseline and 12 week follow-up	

Table D-1. Key question 1 study design details (continued)

Author, Year	Research Objective	Funding Source	Geographic Location, Setting Type, Setting Description	Study Design	Primary Outcomes	Measurement Intervals	Other Notes
Jibaja-Weiss 2003[3]	To evaluate the effectiveness of PF letters and PT letters of prompting communications	Government	US Clinical (In- and Out-Patient) Community health clinics in Houston, Texas, that provide care to underserved and low-income neighborhoods	RCT	Scheduling an appointment for screening and receiving screening services	12 months after receiving the letter	Study was stratified by age and race
Schneider 2001[1]	To examine the effects that differently framed and targeted health messages have on persuading low-income women to obtain screening mammograms	Multiple [Funded by ACS and NCI]	US- though not explicitly stated Community-based settings Community health clinics and public housing	RCT (factorial design)	Self-reported mammography use at 6 months and 12 months [see note]	Baseline, immediate post-test, 6 month follow-up, and 12 month follow-up	Authors cite another study and say "self-reports were correlated reliably with reports in medical records"
Vernon 2008[4] del Junco 2008[5]	To evaluate strategies to increase regular mammography screening	Government	US Community-based settings National Registry of Women Veterans	RCT	Self-reported likelihood of getting a breast cancer screening within 12 and 24 months after exposure to the letter	Baseline, year 1, year 2	Actual survey times were between 6-15 months apart; also called Project HOME

Abbreviations: ACS = American Cancer Society; NCI = National Cancer Institute; PF = personalized form; PT = personalized tailored; RCT = randomized controlled trials; US = United States

Table D-2. Key question 1 sample characteristics, part 1

Author, Year	Groups	Sampling strategy, Unit of Randomization, Process of randomization	Inclusion/ Exclusion Criteria	N Eligible	N Random-ized	N Com-pleters	N analyzed	Other notes
Cox 2001[1]	G1: Control (not abstracted) G2: Gain frame and non-narrative/ statistical G3: Loss frame and non-narrative/ statistical G4: Gain frame and narrative/anecdotal G5: Loss frame and narrative/anecdotal	Convenience Individual Subjects were assigned to one of the five experimental conditions	Women over the age of 50	Overall N=174	Overall N=174	Overall N=174	Overall N=174 G2: N: 29 G3: N: 29 G4: N: 29 G5: N: 29 [see note]	Authors didn't give exact numbers but stated 117 subjects were assigned to 1 of the 4 groups and 57 people where in G1
Elder 2005[2]	G1: Control ("off the shelf" materials covering same modules and content as lay health workers and tailored conditions) (119) G2: Tailored print condition (118) G3: Lay health worker tailored print condition (120)	Random Individual Randomly assigned participants using block randomization	Inclusion: Spanish-language dominant women between 18 and 65 Exclusion: no adult female living in the home; there was no adult female between 18 and 65 years of age; or the target adult female was pregnant, on a special diet for medical reasons, or planning to leave the San Diego area during the study period.	NR	Overall N=357 G1: N: 119 G2: N: 118 G3: N: 120	Overall N=313 G1: N: 107 G2: N: 99 G3: N: 107	Overall N=313 G1: N: 107 G2: N: 99 G3: N: 107	

Table D-2. Key question 1 sample characteristics, part 1 (continued)

Author, Year	Groups	Sampling strategy, Unit of Randomization, Process of randomization	Inclusion/ Exclusion Criteria	N Eligible	N Random-ized	N Com-pleters	N analyzed	Other notes
Jibaja-Weiss 2003[3]	G1: No intervention control (499 for cervical, 239 for breast) G2: PF letters targeted to women age 40 and older (460 for cervical, 239 for breast) G3: PT letter (524 for cervical, 261 for breast)	Computer-generated random numbers Patient 5 waves of recruitment with stratification and then randomization at each stage (except for wave 4 and 5, where they stratified by age only)	Inclusion: had not received mammogram or Pap during past 2 years, had no more than 2 MD visits for acute or chronic illness within past 2 years and had a verifiable mailing address Exclusion: had active breast or cervical cancer, had other serious illnesses, or had received a pap from a local provider	NR	Overall N=1574 G1: N: 499 G2: N: 494 G3: N: 581	Overall N=1483 G1: N: 499 G2: N: 494 G3: N : 581	Overall N=1483 G1: N: 499 G2: N: 460 G3: N: 524	For breast cancer screening outcome, only used those patients age 40+. 17% dropped out after randomization due to back addresses.
Schneider 2001[1]	G1: Gain frame and non-targeted/multicultural G2: Loss frame and non-targeted/multicultural (framing + targeting) G3: Gain frame and targeted toward Latinas G4: Loss frame and targeted toward Latinas	Convenience Sometimes the individual and sometimes the group Women were randomly assigned either individually or at the group level to message framing or not, and targeting or not with a coin flip (first flip is for framing and second for targeting).	Medically underserved women, over the age of 40, who either visit community health clinics or live in public housing	NR	Overall N=752	6 months Overall N=526 12 months Overall N=534	NR	Group sizes not reported

Table D-2. Key question 1 sample characteristics, part 1 (continued)

Author, Year	Groups	Sampling strategy, Unit of Randomization, Process of randomization	Inclusion/ Exclusion Criteria	N Eligible	N Random-ized	N Com-pleters	N analyzed	Other notes
Vernon 2008[4] del Junco 2008[5]	G1: No intervention control (1,840 for 12 months, 754 for 24 months) G2: Targeted (1,857 for 12 months, 825 for 24 months) G3: Targeted and tailored (1,803 for 12 months, 781 for 24 months)	Random number, stratified by sampling round and respondent status Individual STATA was used to generate random numbers for all selection and allocation procedures w/out knowledge of study candidates' characteristics or eligibility survey responses	Inclusion: member of the National registry of women veterans, valid SSN, current mailing address in US or Puerto Rico Exclusion: history of breast CA, physical and mental disability	Overall N=5500 G1: N: 1840 G2: N: 1857 G3: N: 1803	Overall N=5500 G1: N: 1840 G2: N: 1857 G3: N: 1803	Overall N=2681 a (PP) G1: N: 888 G2: N: 907 G3: N: 886	Overall N= 5500 G1: N: 1840 G2: N: 1857 G3: N: 1803	Difficult to tease out the actual numbers for the analyzed groups because they did an ITT, modified ITT, and PP- per protocol analysis.

Abbreviations: G = group; ITT = intent to treat; MD = medical doctor; N = number; NP = __; NR = not reported; PF = personalized form; PP = per protocol; PT = personalized tailored; SSN = social security number; STATA = _.

Table D-3. Key question 1 sample characteristics part 2

Author, Year	Groups	Mean Age	Female	Race, Ethnicity	Median Income	Insured	Education	Health Literacy/ Numeracy	Others Baseline Character-istics (^=signify-cant)	Other Notes
Cox 2001[1]	G1: Control (not abstracted) G2: Gain frame and statistical (framing) G3: Loss frame and statistical (framing) G4: Gain frame and anecdotal (framing + narrative) G5: Loss frame and anecdotal (framing +narrative)	70	100%	67% white NR	NR	NR	90% high school graduates	NR	Marriage status, reported health family member with breast cancer	
Elder 2005[2]	G1: Control ("off the shelf" materials covering same modules and content as lay health workers and tailored conditions) (119) G2: Tailored print condition (118) G3: Lay health worker tailored print condition (120)	39.71	100%	NR 100% Latina	Median NR Income ($ per month) ≤ 1,000: 13.3% 1,001-2,000: 42.3% 2,001-3,000: 29.9% > 3,000: 14.5%	NR	0-6 years: 95, 26.6% Middle school : 89, 24.9% High school: 76, 21.3% Some college: 97, 27.2%	NR	Country of formal education Employment status Self-perceived health^ Marital status Range of dietary variables Total family size Age BMI Waist-hip ratio	

Table D-3. Key question 1 sample characteristics part 2 (continued)

Author, Year	Groups	Mean Age	Female	Race, Ethnicity	Median Income	Insured	Education	Health Literacy/ Numeracy	Others Baseline Character-istics (^=signify-cant)	Other Notes
Jibaja-Weiss 2003[3]	G1: No intervention control (499 for cervical, 239 for breast) G2: PF letters targeted to women age 40 and older (460 for cervical, 239 for breast) G3: PT letter (524 for cervical, 261 for breast)	40.2 [a]	100%	Black: 40.7%[a] Hispanic: 41.7%[a] Non-Hispanic white: 17.6%[a] NR	NR	NR	NR	NR	NR	For breast cancer screening outcome, only used those patients age 40+. 17% dropped out after randomization due to back addresses.
Schneider 2001[1]	G1: Gain frame and multicultural G2: Loss frame and multicultural G3: Gain frame and Latina targeting G4: Loss frame and Latina targeting	56 ± 12 (range 40-91)	100%	White: 27% Black, non-Hispanic: 43% Hispanic: 25% Native American: 1% Asian and Other: 3% NR	$13,500 or less: 62% $13,500 to $18,999: 9% $18,999 or beyond : 10% Rather not report: 18%	NR	Some elementary school: 10% Some middle school (grade 6-8): 14% Some high school: 50% Vocational: 8% Some college: 11% Bachelors or beyond: 6%	NR	Marital status, health ratings, health exam frequency	

Table D-3. Key question 1 sample characteristics part 2 (continued)

Author, Year	Groups	Mean Age	Female	Race, Ethnicity	Median Income	Insured	Education	Health Literacy/ Numeracy	Others Baseline Character-istics (^=signify-cant)	Other Notes
Vernon 2008[4] del Junco 2008[5]	G1: No intervention control (1,840 for 12 months, 754 for 24 months) G2: Targeted (1,857 for 12 months, 825 for 24 months) G3: Targeted and tailored (1,803 for 12 months, 781 for 24 months)	62.38	100%	Black: 7.2% [a] Hispanic: 2.8% [a] Non-Hispanic White: 85.6% [a] Unknown: 4.4% [a] NR	NR-.	NR	High school or less: 14.6% [a] Some college : 44.0% [a] College graduate or higher: 38.4% [a] Unknown: 3.0% [a]	NR	NR	

[a] Calculated by reviewers.

Abbreviations: BMI = body mass index; G = group; NR = not reported; PF = personalized form; PT = personalized tailored

Table D-4. Key question 1 intervention descriptions

Author, Year	Groups	Comparators	Clinical Focus of the Evidence, Source of Information, Theory-Based, Prior Testing	Intervention Format, Delivery Agent, Intensity	Evidence Presen-tation	Message of Intervention	Other Notes
Cox 2001[1]	G1: Control (not abstracted) G2: Gain frame and statistical (framing) G3: Loss frame and statistical (framing) G4: Gain frame and anecdotal (framing + narrative) G5: Loss frame and anecdotal (framing +narrative)	#1: Com - Message framing G2: Statistical evidence with a **gain** frame. G3: Statistical evidence with a **loss** frame. #2: Com – Narratives G4: **Anecdotal** message with a **gain** frame. G5: **Anecdotal** message with a **loss** frame	Mammogram Health education materials by the National Cancer Institute and the American Cancer Society Yes Yes	Paper-based in-person delivery of print advertisement 1 session	G1 & G2: quantitative G3 & G4: qualitative	G2: Message said that "doctors are able to detect tumors at an early, treaTable D-stage, and they are 30% less likely to die from cancer" G3: Message said "doctors are not able to detect tumors at an early, treaTable D-stage, and they are 43% more likely to die of breast cancer." G4: The message had a story about Sara Johnson's ended with "doctors were able to detect her breast tumor at an early, treaTable D-stage, and now Sara can look forward to a long life, watching her grandson, Jeffrey, grow up." G5: The message had a story about Sara Johnson and ended with "doctors were not able to detect her breast tumor at an early, treaTable D-stage, and now Sara may miss out on a long life, watching her grandson, Jeffrey, grow up.”	

Table D-4. Key question 1 intervention descriptions (continued)

Author, Year	Groups	Comparators	Clinical Focus of the Evidence, Source of Information, Theory-Based, Prior Testing	Intervention Format, Delivery Agent, Intensity	Evidence Presen-tation	Message of Intervention	Other Notes
Elder 2005[2]	G1: Control ("off the shelf" materials covering same modules and content as lay health workers and tailored conditions) (119) G2: Tailored print condition (118) G3: Lay health worker tailored print condition (120)	#1: Com- Targeted communication; audience segmentation Targeted newsletters and activities inserts #2: Com - Tailored communication Tailored newsletter and activities insert #3: Com – Narratives Lay health advisor "Promotoras" + tailored newsletter and activities inserts	Reduce dietary fat and increase fiber Unspecified; American heart association NIH; American Dietetic Association, and the American Cancer Society Yes Unclear	G1 & G2: paper-based G3: paper-baesd + in-person G1 & G2: postal G3: Promotoras (characteristics: Spanish-language dominant; naturally empathetic, able to develop rapport and to be neutral and nonjudgmental; perceived as a role model in the community; and interested in helping women change lifestyle behaviors.) G1: one time mailing (probably) G2: 12 weekly mailings G3: 12 weekly mailings of print	NR	G1: Targeted materials were developed for a Latino population and were available in Spanish. Language-appropriate materials that contained information on food purchasing, food preparation, and food consumption were available from the American Heart Association, American Dietetic Association, and the American Cancer Society G2: newsletters provided feedback on the assessment process, as well as an opportunity for personalized goal setting and for dealing with identified barriers. The degree of complexity of the activity in the insert varied by the participant's readiness to change (e.g., acquire information vs. self-monitor). Participants were encouraged to complete the activity on the insert and return the self-addressed stamped card to be entered into a raffle and to receive additional chapters of the story (novela) in the newsletter. There were also magnetic flower petals containing healthy lifestyle messages and eight recipes.	

Table D-4. Key question 1 intervention descriptions (continued)

Author, Year	Groups	Comparators	Clinical Focus of the Evidence, Source of Informa-tion, Theory-Based, Prior Testing	Intervention Format, Delivery Agent, Intensity	Evidence Presen-tation	Message of Intervention	Other Notes
Elder 2005[2] (continued)				materials + 12 weekly home visit or telephone call		G3: Using the skills acquired in the program, as well as their natural ability to provide support and encouragement and their social networking skills, the promotoras worked with individual participants to negotiate behavioral change goals. The promotoras relied primarily on the participant's weekly tailored newsletter to guide discussions and suggest opportunities for skill development.	
Jibaja-Weiss 2003[3]	G1: No intervention control (499 for cervical, 239 for breast) G2: PF letters targeted to women age 40 and older (460 for cervical, 239 for breast) G3: PT letter (524 for cervical, 261 for breast)	#1: Com- Targeted communication; audience segmentation Single page document on clinic letterhead, written at 6th grade level in either English or Spanish and signed by medical director. PF contained generic info on risk factors for breast and cervical cancer, the importance of screening and early detection, and encouragement to schedule a visit for a pelvic examination and pap or clinical breast exam and mammogram.	Breast and cervical cancer screening American Cancer Society Yes Yes	Paper-based mailed letter Single page document mailed to patients #: 1 length: 1 page total time: NR	Quantitative	Included recommendations, breast and cervical cancer risk, and appointment scheduling info.	Primary difference between the PT and PF letters was that the PT letter included personalized breast and cervical cancer risk factors (e.g., Mrs. Smith, you may be at-risk of breast cancer because….) while the PF letter included only standardized phrases about risks. The total possible word counts for the body of each letter were 384 for the PF and 314 for the PT. 80% of the words were in common.

Table D-4. Key question 1 intervention descriptions (continued)

Author, Year	Groups	Comparators	Clinical Focus of the Evidence, Source of Informa-tion, Theory-Based, Prior Testing	Intervention Format, Delivery Agent, Intensity	Evidence Presen-tation	Message of Intervention	Other Notes
Jibaja-Weiss 2003[3] (continued)		The Spanish versions varied only in the use of accepted Spanish-language idioms to approximate the English-language message. #2: Com - Tailored communication Single page document on clinic letterhead, written at 6th grade level in either English or Spanish and signed by medical director, directly addressed to patient. PT contained specific info on 6 risk factors for breast and cervical cancer tailored to the pt: age, race, family history, parity, BMI, and tobacco use, the importance of screening and early detection, and encouragement to schedule a visit for a pelvic exam and pap or CBE and mammogram. Risk factor data were extracted from medical chart.					Based on Health Belief Model.

Table D-4. Key question 1 intervention descriptions (continued)

Author, Year	Groups	Comparators	Clinical Focus of the Evidence, Source of Informa-tion, Theory-Based, Prior Testing	Intervention Format, Delivery Agent, Intensity	Evidence Presen-tation	Message of Intervention	Other Notes
Schneider 2001[1]	G1: Gain frame and multicultural G2: Loss frame and multicultural G3: Gain frame and Latina targeting G4: Loss frame and Latina targeting Group sizes not reported	#1: Com - Message framing Gain-framed video: emphasized the benefits of getting a mammogram #2: Com - Message framing Loss-framed video: emphasized the costs of NOT getting a mammogram #3: Com- Targeted communication; audience segmentation Gain-framed AND targeted video- where video has benefits and shows >60% of photos are race specific and text is race specific (Anglo, Black, Latina) and 40% of narrative was framed #4: Com- Loss-framing and targeted video	Breast cancer screening NCI and ACS information Yes Yes	Video Research assistant of unknown type turns video on. #: 1 length: 10 minutes total time: 10 minutes	Combined (multi-media)	Various motivational messages about mammography and breast cancer screening.	

Table D-4. Key question 1 intervention descriptions (continued)

Author, Year	Groups	Comparators	Clinical Focus of the Evidence, Source of Informa-tion, Theory-Based, Prior Testing	Intervention Format, Delivery Agent, Intensity	Evidence Presen-tation	Message of Intervention	Other Notes
Vernon 2008[4] del Junco 2008[5]	G1: No intervention control (1,840 for 12 months, 754 for 24 months) G2: Targeted (1,857 for 12 months, 825 for 24 months) G3: Targeted and tailored (1,803 for 12 months, 781 for 24 months)	#1: Com- Targeted communication; audience segmentation a folder containing 1) a set of 4 educational booklets, 2) a letter for the women to use to discuss mammography with her PCP and 3) a pamphlet about mammography screening services at the VA #2: Com - Tailored communication Received both Targeted info and Tailored component: A 4 page letter with messages addressed 1)woman's stage of change 2) feedback regarding her decisonal balance 3) graphical illustrations of her objective and perceived risks for breast CA and messages to reconcile the two, 4) feedback on her self-	Breast cancer screening ACS Yes Unclear	Paper-based Postal #: 2 length: NR total time: over 3.25 yrs	combined	Educational booklet, local info on services, perceived and actual risk factors, motivational info,	Some of the constructs about stages of change were used prior, but not sure about actual interventions, doesn't say about intervention

Table D-4. Key question 1 intervention descriptions (continued)

Author, Year	Groups	Comparators	Clinical Focus of the Evidence, Source of Informa-tion, Theory-Based, Prior Testing	Intervention Format, Delivery Agent, Intensity	Evidence Presen-tation	Message of Intervention	Other Notes
Vernon 2008[4] del Junco 2008[5] (continued)		efficacy, 5) review of her use of the process of change and activities she could appropriate for her stage of change, 6) reminder about her next mammography due date					

Abbreviations: ACS = American Cancer Society; BMI = body mass index; CA = cancer; CBE = __; Com = communication; G = group; NCI = National Cancer Institute; NIH = National Institute of Health; NR = not reported; PCP = primary care physician; PF = personalized form; PT = personalized tailored; VA = Veteran's Administration; vs. = versus.

Table D-5. Key question 1 first outcome

Author, Year	Groups	Outcome #1, Exact Measure Used	Timing of Measurement, Data Source	N analyzed for This Outcome	Results by Group	Differences in Groups	Statistical Methods Used, Covariates Controlled for in Analysis
Cox 2001[1]	G1: Control (not abstracted) G2: Gain frame and statistical (framing) G3: Loss frame and statistical (framing) G4: Gain frame and anecdotal (framing + narrative) G5: Loss frame and anecdotal (framing +narrative)	Behavioral intentions to use or apply the evidence Perceived likelihood of having a mammogram after seeing advertisement. Higher numbers indicate greater perceived likelihood of getting a mammogram	Immediate posttest Self-report	174 overall G2: 29 G3: 29 G4: 29 G5: 29	Mean likelihood (7-point Likert scale where a higher number means greater likelihood): G2: 5.48 G3: 4.37 G4: 4.07 G5: 5.54	Significant interaction effect: F (1,103): 10.87, p=.001 G4 vs. G5: 1.47 [a] (p<.01) G2 vs. G3: 1.11 [a] (p=.06) G2 vs G4: 1.41 [a] (ns) G2 vs G5: 1.17 [a] (ns) G3 vs G4: 0.3 [a] (ns) G3 vs G5: 1.17 [a] (p<.01)	ANOVA NR
Elder 2005[2]	G1: Control ("off the shelf" materials covering same modules and content as lay health workers and tailored conditions) (119) G2: Tailored print condition (118) G3: Lay health worker tailored print condition (120)	Clinical outcomes % calories from fat	Baseline and 12 week follow-up Self-report face-to-face interview	Baseline N=357 G1: 119 G2: 118 G3: 120 Follow-up N=313 G1: 107 G2: 99 G3: 107	Percentage at baseline minus percentage at 12 weeks G1: 31.5-30.0=1.5 G2: 31.0-30.4=0.6 G3: 31.5-29.3= 2.2	Difference of differences between G2 vs. G1: -0.9 [a] (favoring G1=fewer calories from fat) Differences among the 3 groups at 12 weeks controlling for baseline level not significant F=0.81, p=0.45	Tukey-Kramer multiple comparison test Baseline measure

Table D-5. Key question 1 first outcome (continued)

Author, Year	Groups	Outcome #1, Exact Measure Used	Timing of Measurement, Data Source	N analyzed for This Outcome	Results by Group	Differences in Groups	Statistical Methods Used, Covariates Controlled for in Analysis
Jibaja-Weiss 2003[3]	G1: No intervention control (499 for cervical, 239 for breast) G2: PF letters targeted to women age 40 and older (460 for cervical, 239 for breast) G3: PT letter (524 for cervical, 261 for breast)	Clinical outcomes Scheduling a pap appointment- EHR record of appointments made and receiving a Pap- EHR record of completed visit	Within 12 months after letter was sent Medical record	Overall N=1483	Scheduled- p<0.001 Percentage G1: 44.7% G2: 53.3% G3: 39.7% Received- p<0.001 Percentage G1: 39.9% G2: 43.9% G3: 23.7%	Scheduling Difference G2-G1: 8.6% [a] Scheduling Difference G3-G2: -13.6% [a] Scheduling Difference G3-G1: -5% [a] Screened Difference G2-G1: 4% [a] Screened Difference G3-G2: -20.2% [a] Screened Difference G3-G1: -16.2% [a]	Chi-squared None

Table D-5. Key question 1 first outcome (continued)

Author, Year	Groups	Outcome #1, Exact Measure Used	Timing of Measurement, Data Source	N analyzed for This Outcome	Results by Group	Differences in Groups	Statistical Methods Used, Covariates Controlled for in Analysis
Schneider 2001[1]	G1: Gain frame and multicultural G2: Loss frame and multicultural G3: Gain frame and Latina targeting G4: Loss frame and Latina targeting Group sizes not reported	Clinical outcomes Self-reported likelihood of getting a mammogram in the past 12 months within 6 months after seeing video, comparing "loss" to "gain" frames holding type of targeting constant; and looking at their interactive effect	6 month call or postcard Self-report	Overall N=752	Percentage Overall= 41% G1: 36% G2: 50% G3: 41% G4: 36%	G2 vs. G1: 14% [a]=OR =1.81 (p<0.01) (favoring G2) G4 vs. G3: -5%[a]=OR=1.22 (p=0.10) (favoring G3) Using hierarchical logistic regression, controlling for past year's use (6 months after exposure): Framing x Targeting interaction: Chi-square: 5.15, p<0.05 OR [CIs]: Past year's Mammography use: 1.44 [0.98, 2.11] Loss framing: 1.27 [0.78, 2.08] Targeting: 1.20 [0.72,1.99] Frame x Target: 2.27 [1.12, 4.63]	Absolute differences None

Table D-5. Key question 1 first outcome (continued)

Author, Year	Groups	Outcome #1, Exact Measure Used	Timing of Measurement, Data Source	N analyzed for This Outcome	Results by Group	Differences in Groups	Statistical Methods Used, Covariates Controlled for in Analysis
Vernon 2008[4] del Junco 2008[5]	G1: No intervention control (1,840 for 12 months, 754 for 24 months) G2: Targeted (1,857 for 12 months, 825 for 24 months) G3: Targeted and tailored (1,803 for 12 months, 781 for 24 months)	Clinical outcomes Self-reported likelihood of getting a breast cancer screening within 12 months after exposure to the letter	Year 1 Self-report	Overall N=5500 G1: N: 1840 G2: N: 1857 G3: N: 1803	Crude Incidence using ITT analysis: G1: 44.7% G2: 46.9% G3: 46.0%	ITT difference G3 vs. G2: -0.9% [a] (favoring G2) Chi-square: 1.70 2 d.f . p= 0.427 Cox proportional hazard rate ratio [CI] using ITT: Differences were not significant. G1: 1.00 G2: 1.07 [0.97,1.18] G3: 1.05 [0.95,1.15]	Chi-squared None

Abbreviations: ANOVA = analysis of variance; CI = confidence interval; d.f. = degrees of freedom; EHR = electronic health record; G = group; ITT = intention to treat; NR = not reported; ns = not significant; OR = odds ratio

Table D-6. Key question 1 second outcome

Author, Year	Groups	Outcome #2, Exact Measure Used	Timing of Measurement, Data Source	N Analyzed for This Outcome	Results by Group	Differences in Groups	Statistical Methods Used, Covariates Controlled for in Analysis
Cox 2001[1]	G1: Control (not abstracted) G2: Gain frame and statistical (framing) G3: Loss frame and statistical (framing) G4: Gain frame and anecdotal (framing + narrative) G5: Loss frame and anecdotal (framing +narrative)	Knowledge about the evidence Risk factor knowledge: (e.g., can develop breast cancer without symptoms; can develop breast cancer without a family history of breast cancer). Higher numbers indicate greater risk factor knowledge	Immediate post test Self-report	174 overall G2: 29 G3: 29 G4: 29 G5: 29	Means G2: -0.20 G3: -0.10 G4: 0.09 G5: 0.01	No main effects, no interaction effects. G2 vs. G3: 0.10 [a] G2 vs. G4: 0.29 [a] G2 vs. G5: 0.21 [a] G3 vs. G4: 0.19 [a] G3 vs. G5: 0.11 [a] G4 vs. G5: 0.08 [a] (all ns)	ANOVA NR
Elder 2005[2]	G1: Control ("off the shelf" materials covering same modules and content as lay health workers and tailored conditions) (119) G2: Tailored print condition (118) G3: Lay health worker tailored print condition (120)	Clinical outcomes Total dietary fiber (g)	Baseline and 12 week follow-up Self-report face-to-face interview	Follow-up N=313 G1: 107 G2: 99 G3: 107	Adjusted mean, in grams at baseline minus grams at 12 weeks G1: 16.5-15.6=0.9 G2: 17.2-17.2=0.0 G3: 17.2-16.1=1.1	Difference of the differences between G2 vs.G1: -0.09 [a] (favoring G1=more grams of fiber) Differences among the 3 groups at 12 weeks controlling for baseline level not significant F=1.61, p=0.20, not significant	Tukey-Kramer multiple comparison test Baseline mean

Table D-6. Key question 1 second outcome (continued)

Author, Year	Groups	Outcome #2, Exact Measure Used	Timing of Measurement, Data Source	N Analyzed for This Outcome	Results by Group	Differences in Groups	Statistical Methods Used, Covariates Controlled for in Analysis
Jibaja-Weiss 2003[3]	G1: No intervention control (499 for cervical, 239 for breast) G2: PF letters targeted to women age 40 and older (460 for cervical, 239 for breast) G3: PT letter (524 for cervical, 261 for breast)	Clinical outcomes Scheduling a breast cancer screening appointment- EHR record of appointments made and receiving a mammogram- EHR record of completed mammogram	12 months after letter was sent Medical chart	N=739	Scheduled- p<0.001 Percentage G1: 53.3% G2: 65.7% G3: 50.2% Received- p<0.001 Percentage G1: 20.7% G2: 30.5% G3: 13.0%	Scheduling Difference G2-G1: 12.4% [a] Scheduling Difference G3-G2: -15.5% [a] Scheduling Difference G3-G1: -3.1% [a] Screened Difference G2-G1: 9.8% [a] Screened Difference G3-G2: -17.5% [a] Screened Difference G3-G1: -7.7% [a]	Chi-squared None

Table D-6. Key question 1 second outcome (continued)

Author, Year	Groups	Outcome #2, Exact Measure Used	Timing of Measurement, Data Source	N Analyzed for This Outcome	Results by Group	Differences in Groups	Statistical Methods Used, Covariates Controlled for in Analysis
Schneider 2001[l]	G1: Gain frame and multicultural G2: Loss frame and multicultural G3: Gain frame and Latina targeting G4: Loss frame and Latina targeting Group sizes not reported	Clinical outcomes Self-reported likelihood of getting a mammogram in the past 12 months within 12 months after seeing video, comparing "loss" to "gain" frames holding type of targeting constant; and looking at their interactive effect	12 month call or postcard only to people who didn't get mammogram at 6 months Self-report	NR	Percentage Overall: 57% G1: 55% G2: 61% G3: 57% G4: 54%	G2 vs. G1: 6% [a] Not significant, p value not reported (favoring G2) G4 vs. G3: -3% [a] Not significant, p value not reported (favoring G3) Using hierarchical logistic regression, controlling for past year's use (12 months after exposure): Framing x Targeting interaction: Chi-square=1.65, ns OR [CIs]: Past year's Mammography use: 2.93 [2.05, 4.18], p<0.01 Loss framing: 1.18 [0.74, 1.89] Targeting: 1.05 [0.65, 1.70] Frame x Target: 1.56 [0.79, 3.08]	Absolute differences None

Table D-6. Key question 1 second outcome (continued)

Author, Year	Groups	Outcome #2, Exact Measure Used	Timing of Measurement, Data Source	N Analyzed for This Outcome	Results by Group	Differences in Groups	Statistical Methods Used, Covariates Controlled for in Analysis
Vernon 2008[4] del Junco 2008[5]	G1: No intervention control (1,840 for 12 months, 754 for 24 months) G2: Targeted (1,857 for 12 months, 825 for 24 months) G3: Targeted and tailored (1,803 for 12 months, 781 for 24 months)	Clinical outcomes Self-reported likelihood of getting a breast cancer screening within 24 months after exposure to the letter	Year 2 Self-report	Overall N= 2360 G1: N: 754 G2: N: 825 G3: N: 781	Crude Incidence using ITT analysis: G1: 22.0% G2: 24.8% G3: 24.8%	ITT difference G3 vs. G2: 0.0% a Chi-square: 5.17 2 d.f. p=0.075 (G2 and G3 are equal) Cox proportional hazard rate ratio [CI] using modified ITT: Differences not significant. G1: 1.00 G2: 0.99 [0.86 to 1.13] G3: 1.05 [0.91 to 1.20]	Chi-squared None

Abbreviations: ANOVA = analysis of variance; CI = confidence interval; d.f. = degrees of freedom; g = gram; G = group; HER = electronic health record; ITT = intention to treat; N = number; NR = not reported; ns = not significant; OR = odds ratio; PF = personalized form; vs. = versus

Table D-7. Key question 1 third outcome

Author, Year	Groups	Outcome #3, Exact Measure Used	Timing of Measurement, Data Source	N Analyzed for This Outcome	Results by Group	Differences in Groups	Statistical Methods Used, Covariates Controlled for in Analysis
Cox 2001[1]	G1: Control (not abstracted) G2: Gain frame and statistical (framing) G3: Loss frame and statistical (framing) G4: Gain frame and anecdotal (framing + narrative) G5: Loss frame and anecdotal (framing +narrative)	NA	NA	NA	NA	NA	NA
Jibaja-Weiss 2003[3]	G1: No intervention control (499 for cervical, 239 for breast) G2: Personalized form (PF) letters targeted to women age 40 and older (460 for cervical, 239 for breast) G3: Personalized tailored (PT) letter (524 for cervical, 261 for breast)	NA	NA	NA	NA	NA	NA
Vernon 2008[4] del Junco 2008[5]	G1: No intervention control (1,840 for 12 months, 754 for 24 months) G2: Targeted (1,857 for 12 months, 825 for 24 months) G3: Targeted and tailored (1,803 for 12 months, 781 for 24 months)	NA	NA	NA	NA	NA	NA

Table D-7. Key question 1 third outcome

Author, Year	Groups	Outcome #3, Exact Measure Used	Timing of Measurement, Data Source	N Analyzed for This Outcome	Results by Group	Differences in Groups	Statistical Methods Used, Covariates Controlled for in Analysis
Elder 2005[2]	G1: Control ("off the shelf" materials covering same modules and content as lay health workers and tailored conditions) (119) G2: Tailored print condition (118) G3: Lay health worker tailored print condition (120)	NA	NA	NA	NA	NA	NA

[a] Calculated by reviewers.

Abbreviations: Com=comparator; d.f. = degrees of freedom; NA = not applicable; ns = nonsignificant; US = United States

References

1. Cox D, Cox AD. Communicating the consequences of early detection: The role of evidence and framing. J Market. 2001;65(3):91-103.
2. Elder JP, Ayala GX, Campbell NR, et al. Long-term effects of a communication intervention for Spanish-dominant Latinas. Am J Prev Med. 2006 Aug;31(2):159-66. PMID: 16829333.
3. Jibaja-Weiss ML, Volk RJ, Kingery P, et al. Tailored messages for breast and cervical cancer screening of low-income and minority women using medical records data. Patient Educ Couns. 2003 Jun;50(2):123-32. PMID: 12781927.
4. Vernon SW, del Junco DJ, Tiro JA, et al. Promoting regular mammography screening II. Results from a randomized controlled trial in US women veterans. J Natl Cancer Inst. 2008 Mar 5;100(5):347-58. PMID: 18314474.
5. del Junco DJ, Vernon SW, Coan SP, et al. Promoting regular mammography screening I. A systematic assessment of validity in a randomized trial. J Natl Cancer Inst. 2008 Mar 5;100(5):333-46. PMID: 18314473.

Appendix E. Evidence Tables for Key Question 2

Table E-1. Key question 2 study design details

Author, Year	Research objective	Funding Source	Geographic location, Setting type, Setting Description	Study design	Primary Outcomes	Measurement intervals	Other Notes
Bahrami et al., 2004[1]	The aim of this study was to evaluate the effectiveness of different implementation strategies for evidence-based clinical guidelines using SIGN as a model	Government	Scotland Clinical (In- and Out-Patient) dental practices	cRCT	Compliance with guideline	Two 4 month periods pre and post intervention	
Banait et al., 2003[2]	To test the effectiveness of 'educational outreach' as a strategy for facilitating the uptake of dyspepsia management guidelines in primary care.	Unspecified	England Clinical (In- and Out-Patient) General practices in the Salfrod & Trafford Health authority catchment area in NW England	RCT	Appropriateness of referrals for open access endoscopy Findings at open access endoscopy Prescribing costs	7 months pre and post intervention	
Beaulieu et al., 2004[3]	To study the effects of guideline dissemination on physicians' prescribing practices for the treatment of sTable E-angina pectoris.	Government	Canada Clinical (In- and Out-Patient) Practicing physicians in urban, suburban, and rural parts of Quebec, Canada	RCT	Prescription of 3 cardiovascular medications in 1999	6 month follow-up	
Becker et al., 2008[4]	To improve quality of care for patients with low back pain (LBP) a multifaceted general practitioner education alone and in combination with motivational counseling by practice nurses has been implemented in German general practices.	Government	Germany Clinical (In- and Out-Patient) General practices in semi rural German regions	cRCT	Functional capacity	Baseline and 6 and 12 month follow-up	Physical activity during 1 week before interview, days in pain and days of sick leave during 6 months follow-up, quality of life

Table E-1. Key question 2 study design details (continued)

Author, Year	Research objective	Funding Source	Geographic location, Setting type, Setting Description	Study design	Primary Outcomes	Measurement intervals	Other Notes
Bekkering et al., 2005[5, #4960]	To evaluate the effect on the process of care of an active strategy to implement clinical guidelines on physiotherapy for low back pain.	Unspecified	Netherlands Clinical (In- and Out-Patient) Private practices in the center of the Netherlands	cRCT	Adherence to guideline; Physical functioning, pain, and sick leave	Baseline and followup; Baseline, 6, 12, 26, and 52 weeks after baseline	
Bishop et al., 2006[6]	The goal of this study was to determine whether or not providing both family physicians and their patients with information about clinical practice guidelines in a direct and individualized manner would increase guideline concordance.	Workers Compensation Board of British Columbia	Canada Clinical (In- and Out-Patient) NR (possibly more information in previous article)	RCT	Guideline-concordant and -discordant treatment advice and procedures	0-4 weeks, 5-12 weeks, >12 weeks	
Campbell et al., 2004[7]	Compare the effectiveness of 2 strategies to promote colorectal cancer preventive behaviors among 587 African American members of 12 rural North Carolina churches	Government	United States Community-based settings NR (possibly more information in previous article)	fRCT	Diet – fruit and vegeTable E-consumption Physical activity CRC screening	Baseline and 1 year follow-up	
Carney et al., 2005[8]	Tested the impact of two interventions on a population-based sample of NH women who were not receiving routine mammography to determine if adherence to screening could be improved.	ACS and NCI funding	United States Other NR	Randomized trial	Adherence to mammography screening	Baseline and 12 months later	

Table E-1. Key question 2 study design details (continued)

Author, Year	Research objective	Funding Source	Geographic location, Setting type, Setting Description	Study design	Primary Outcomes	Measurement intervals	Other Notes
Christakis et al., 2006[9]	To test the hypothesis that parental activation could occur through directed use of an Internet site before a well-child visit and that this activation would promote the discussion of evidence-based prevention topics with providers and would result in increased parental and physician adoption of preventive measures.	Government	United States Clinical (In- and Out-Patient) 4 nonteaching clinics in the University of Washington Physician Network	fRCT	Discussion of MyHealthy Child Topics Implementation of MyHealthy Child Topics	Baseline and then up to 365 days after baseline.	
Davis et al., 2004[10]	To determine the effectiveness of two dissemination and implementation strategies to implement a national guideline for epilepsy management in primary care settings.	Unspecified	Scotland Clinical (In- and Out-Patient) General practices in Tayside (UK)	cRCT	SF-36 general health-related quality-of-life instrument	baseline and 12 months later	
Eaton et al., 2011[11]	To determine whether an intervention based on patient activation and a physician decision support tool was more effective than usual care for improving adherence to National Cholesterol Education Program guidelines	Unspecified	U.S. Clinical (In- and Out-Patient) Primary care practices	cRCT	Percentage of patients screened for hyperlipidemia and treated to their low-density lipoprotein (LDL) and non–high-density lipoprotein (HDL) cholesterol goals	1 year	

Table E-1. Key question 2 study design details (continued)

Author, Year	Research objective	Funding Source	Geographic location, Setting type, Setting Description	Study design	Primary Outcomes	Measurement intervals	Other Notes
Elder et al., 2005[12]	The present study examined two innovative lifestyle behavior change approaches to reduce dietary fat and to increase fiber. Analyses emphasized (a) whether personalized counseling via promotora plus tailored print materials used in an interactive format were more effective than tailored materials develivered in a distance learning format, and (b) whether these two innovations were more effective than standard off-the-shelf materials TARGETED (culturally) to a Latino population (controls).	Government	United States Community-based settings Setting comprised 2 contiguous metropolitan statistical areas (MSA) within San Diego County. Latinos comprise 53% and 36%, respectively, of the population within these two areas.	RCT	Percent calories from fat Number of daily grams of fiber	Baseline and 12 week follow-up	
Feldstein et al., 2006[13]	To evaluate methods to increase guideline-recommended osteoporosis care postfracture	Pharmaceutical	US Clinical (In- and Out-Patient) Non-profit, group-model HMO in the Pacific Northwest with about 454,000 members	Random-ized trial	Proportion of study population who received a pharmacological treatment or a bone mineral density measurement within 6 months after the intervention	6 months post intervention	

Table E-1. Key question 2 study design details (continued)

Author, Year	Research objective	Funding Source	Geographic location, Setting type, Setting Description	Study design	Primary Outcomes	Measurement intervals	Other Notes
Gattellari et al., 2005[14]	To compare the impact of 3 information resources about PSA screening and determine the extent to which man's preferences for involvement in decision-making change the impact of the resources.	Unspecified	Australia Community-based settings Community dwelling sample of men in 29 contiguous postcodes in Sydney, Australia, found in the white-page telephone directory.	Random-ized trial	Knowledge about prostate cancer	21 days median length between pre-test and post-test	
Hagmolen et al., 2008[15]	Investigates whether written treatment advice to the GP (via the introduction of a national guideline)--based on symptoms, medication use, lung function, and the severity of AHR--results in an improvement in children's asthma after one year.	Pharmaceutical	Netherlands Clinical (In- and Out-Patient) Centralized health care organization with 18 health care centers	cRCT	Change in AHR in children after one year.	Baseline and 1 year later	
Jain et al., 2006[16]	To compare the effectiveness of active to passive dissemination of the Canadian clinical practice guidelines (CPGs) for nutrition support for the mechanically ventilated critically ill adult patient.	Government	Canada Other Centralized health care organization with 18 health care centers	cRCT	Nutritional adequacy of enteral nutrition	Baseline and 12 months later	Both academic and community settings

Table E-1. Key question 2 study design details (continued)

Author, Year	Research objective	Funding Source	Geographic location, Setting type, Setting Description	Study design	Primary Outcomes	Measurement intervals	Other Notes
Jousimaa et al., 2002[17]	To compare the effects of computerized and paper-based versions of guidelines on recently qualified physicians' consultation practices	Private trust, foundation, professional organization	Finland Clinical (In- and Out-Patient) Primary care clinics	cRCT	Physicians' compliance with guideline recommendations about laboratory, radiologic, physical, and other examinations; procedures; nonpharmacologic and pharmacologic treatments; physiotherapy; and referrals	One month post-intervention	
Junghans et al., 2007[18]	Determine the effect of patient-specific ratings vs. conventional guidelines on appropriate investigation (test ordering) of angina	Government	England Clinical (In- and Out-Patient) Clinical practice in Scotland and England	RCT	Agreement of physicians recommendations with those made by 2 independent expert panels.	Immediate posttest	
Kennedy et al., 2003[19]	To develop decision aids to provide evidence-based information and formal preference elicitation for women with menorrhagia; and to evaluate their effects on patient outcomes, patient management and cost-effectiveness.	Government	England Clinical (In- and Out-Patient) Six hospitals in south-west England	RCT	health status	Baseline (6-weeks preconsultation), immediate post-consultation, 6, 12, and 24 months post-consultation. NOTE: 6-month and 12-month data merged together to form a short-term follow-up dataset.	

Table E-1. Key question 2 study design details (continued)

Author, Year	Research objective	Funding Source	Geographic location, Setting type, Setting Description	Study design	Primary Outcomes	Measurement intervals	Other Notes
King et al., 2007[20]	This study determined the 6- and 12-month effectiveness of telephone interventions delivered by health educators or by an automated computer system in promoting physical activity.	Government	United States Community-based settings Doesn't provide details about community	RCT	Physical activity	Baseline, 6, and 12months	
Laprise et al., 2009[21]	The main objective was to determine if the PER intervention delivered by the nurse after the CME activity increased GPs' adherence to CPGs' recommendations	Pharmaceutical	Canada Community-based settings GPs in 5 regions of Quebec	cRCT	GP adherence to CPG recommend-ations	Baseline, 6 month followup	
Lien et al., 2007,[22] Svetkey et al., 2003,[23] Young et al., 2009[24]	This article describes the impact of PREMIER behavioral interventions on BP, lipids, and insulin resistance in subgroups defined by the presence or absence of MetSyn.	Government	United States Clinical (In- and Out-Patient) Participating institutions include the NHLBI Project Office, the Coordinating Center and four clinical centers (Johns Hopkins University; Pennington Biomedical Research Center; Duke University Medical Center; and a clinical center also located at the Kaiser Permanente Center for Health Research.	RCT	6 month change in systolic blood pressure. Weight reduction, improved fitness, lower sodium intake, meet health goals.	Baseline, 3 month, 6 month, 12 month, 18 month	Different data collected at different time points

Table E-1. Key question 2 study design details (continued)

Author, Year	Research objective	Funding Source	Geographic location, Setting type, Setting Description	Study design	Primary Outcomes	Measurement intervals	Other Notes
Marcus et al., 2009[25]	To determine whether one of 2 delivery channels (telephone and print) was more effective in promoting physical activity	Government	US Other NR	RCT	Physical activity recall, fitness data, stage of change	Baseline, 6 months and 12 months	
Maxwell et al., 2010[26]	To develop a Multi-component intervention that would increase colorectal cancer screening among an Asian American population	Foundation or non-profit	USA Community-based settings Small groups met at community organizations in California (may be a smaller area but no details)	cRCT	Self-reported CRC Screening rates	Baseline, 6 month follow-up	
Murtaugh et al., 2005[27]	To test the effectiveness of two interventions designed to improve the adoption of evidence-based practices by home health nurses caring for heart failure (HF) patients.	Government	USA- though not explicitly stated Clinical (In- and Out-Patient) RN clinical visits as part of a large, urban, nonprofit home health agency	RCT	Practice of Evidence-based care (recording key assessment items and instructions to patients, instructing patients with key educational elements)	RN note within 45 days after initial home health RN assessment	Authors talk about US Medicare rules but never specify study in US, though authors are from NYC, infer it was done in NYC; and it looks like the authors abstracted 1 note/pt, but not sure.
Paradis et al., 2011[28]	To test the feasibility, impact, and acceptance of incorporating a DVD of newborn anticipatory guidance into routine wellchild care.	Professional organization and a foundation	U.S. Academic health care institutions Large hospital-based primary care pediatric practice	Random-ized trial	Parent knowledge of infant development; self-efficacy with infant care skills; problem-solving competence	Baseline, 2 weeks and 2 months post-intervention	

Table E-1. Key question 2 study design details (continued)

Author, Year	Research objective	Funding Source	Geographic location, Setting type, Setting Description	Study design	Primary Outcomes	Measurement intervals	Other Notes
Partin et al., 2004[29]	The primary objective of this study was to assess the relative effectiveness of the video and a mailed pamphlet intervention for increasing patient CaP screening knowledge and decision-making participation.	Government	United States Clinical (In- and Out-Patient) Veteran Affairs medical facilities in the Midwest	RCT	Prostate cancer screening knowledge	1 week post	
Rahme et al., 2005[30]	Examined whether a continuing medical education intervention increased general practitioners' ability to select the proper pharmacological treatment for patients with osteoarthritis.	Pharmaceutical	Canada Clinical (In- and Out-Patient) General practitioners from 8 small towns of relatively small population sizes (30K-50K)	cRCT	Dispensed prescriptions (prescription adequacy)	135 to 1 day pre-intervention; 1 to 136 days post intervention	
Rebbeck et al., 2006[31]	To evaluate the effect of an active dissemination strategy that included education by opinion leaders compared with a passive dissemination strategy that consisted of dissemination of the guidelines only via mail	Government	Australia Clinical (In- and Out-Patient) Physiotherapy clinics	cRCT	Patient outcomes: disability, disability due to whiplash, change in symptoms (global perceived effort), patient satisfaction with care Physiotherapist outcomes: knowledge, clinical practice based on guidelines, satisfaction with guidelines	Patient outcomes: Baseline, 1.5 months, 3 months, 6 months, and 12 months. Physiotherapist outcomes: baseline and 12 months	

Table E-1. Key question 2 study design details (continued)

Author, Year	Research objective	Funding Source	Geographic location, Setting type, Setting Description	Study design	Primary Outcomes	Measurement intervals	Other Notes
Rimer et al., 2001[32]	To compare tailored print materials +/-tailored telephone counseling to usual care for promoting mammography screening	Government	USA Clinical (In- and Out-Patient) Blue Cross Blue Shield in North Carolina	RCT	Main outcome was accuracy of risk perception; knowledge, and adherence to yearly screening mammography	Yearly	
Simon et al., 2005[33]	To compare group versus individual academic detailing to increase diuretic or beta blocker use in hypertension.	Government	U.S. Clinical (In- and Out-Patient) Geographically separated HMO practices	cRCT	Change in guideline adherence (the proportion of patients with incident hypertension receiving a diuretic ir beta blocker)	Baseline, 1-year follow-up, 2-year follow-up	
Soler et al., 2010[34]	To determine if dissemination of guidelines plus training and use of a portable-device to perform spirometry tests led to improved diagnosis and categorization of COPD, improved management of COPD, and reduction in other diagnostic interventions	Multiple	Spain Clinical (In- and Out-Patient) General practices in Spain	RCT	Improved diagnosis, severity classification and management of COPD patients in primary care.	Data collected starting 45 days post training and continued for 3 months	
Sullivan et al., 2010[35]	To determine if an interactive web-based training focusing on shared decision-making for chronic opioid therapy improves knowledge and competence compared with exposure to practice guidelines.	Pharmaceutical	U.S. Other All academic hospitals except for one non-academic, urban hospital	RCT	Residents' knowledge, competence, and satisfaction with managing opioids for CNCP	Baseline, immediate posttest, 60-day posttest	

Table E-1. Key question 2 study design details (continued)

Author, Year	Research objective	Funding Source	Geographic location, Setting type, Setting Description	Study design	Primary Outcomes	Measurement intervals	Other Notes
Watson et al., 2002[36]	to compare the effectiveness and efficiency of two guideline dissemination strategies in community pharmacy settings	Government	Scotland Community-based settings All eligible pharmacies (n=121) in the Grampian region of Scotland	cRCT	1. proportion of visits resulting in an appropriate sale or non-sale of an anti-fungal product (based upon the guideline recommendations) 2.pharmacists knowledge of the treatment of vaginal candidiasis	Baseline, immediate posttest timing of posttest not specified	
Wetter et al., 2006[37]	To evaluate 1) paid media approaches for increasing the utilization of the CIS Spanish-language smoking cessation counseling services, and 2) the efficacy of an enhanced counseling intervention for helping Spanish-speaking smokers quit.	Government	U.S. Community-based settings At home over the phone	RCT	Smoking abstinence	Baseline, 5-week follow-up, 12-week follow-up	
Wolters et al., 2005[38]	To determine the effect of a distance learning program on general practice management of men with lower urinary tract symptoms	Academic	Netherlands Clinical (In- and Out-Patient) Clinic in the Netherlands	cRCT	# of PSA requests Medication prescribed Referral rate to a urologist	Baseline Up to 1 year post intervention	

Table E-1. Key question 2 study design details (continued)

Author, Year	Research objective	Funding Source	Geographic location, Setting type, Setting Description	Study design	Primary Outcomes	Measurement intervals	Other Notes
Wright et al., 2008[39]	To examine the effectiveness of local and expert opinion leaders on improving lymph node assessment for patients with stage II colon cancer.	Government, foundation, and academic	Canada Clinical (In- and Out-Patient) Academic and non-academic hospitals in Ontario	cRCT	1) Mean # of lymph nodes assessed in patients with stage II colon cancer; 2) the proportion of cases staged with a minimum of 12 lymph nodes	360 days before intervention, 360 days after intervention	

Abbreviations: CaP = Cancer of the Prostate; CIS = Computer Information Service; CNCP = chronic non-cancer pain; COPD = Chronic obstructive pulmonary disease; cRCT = clustered randomized controlled trial; HMO = health maintenance organization; PSA = Prostate-specific antigen; RCT = randomized controlled trial; U.S. = United States; USA = United States of America.

Table E-2. Key question 2 sample characteristics, part 1

Author, Year	Groups	Sampling Strategy, Unit of Randomization, Process of Randomization	Inclusion/ Exclusion Criteria	N Eligible	N Randomized	N Completers	N Analyzed	Other Notes
Bahrami et al., 2004[1]	G1: Mailed guideline (increase reach) G2: Guideline + AF (not abstracted) G3: CAL (increase ability) G4: CAL + AF (not abstracted)	Random (convenience sample due to volunteering to participate, then randomized) Practice Computer generation of random # sequence performed by an statistician independent of the research	Practice associated with the Scottish Dental Practice Board. Patient has to be 16-24 who attended their dental surgery over two, four-month periods in 1999 pre- and 2000 post-intervention	N: 565 practices	Dentists 51 G1: 12 G3: 13	N: 47 G1: 11 G3: 11	N: 47 G1: 11 G3: 11	
Banait et al., 2003[2]	G1: Mailed guidelines (increase reach) G2: Educational outreach (Multi-component)	Convenience Practice location Minimization. The criteria used for minimization were practice size, fundholding status, previous expenditure on acid-suppressing drugs, and previous involvement in a local guideline initiative.	All general practices in Greater Manchester	n = 115	n = 114 G1: 56 G2: 57	G1: 33 G2: 56	G1: 57 (analyzed using ITT) G2: 56	

Table E-2. Key question 2 sample characteristics, part 1 (continued)

Author, Year	Groups	Sampling Strategy, Unit of Randomization, Process of Randomization	Inclusion/ Exclusion Criteria	N Eligible	N Randomized	N Completers	N Analyzed	Other Notes
Beaulieu et al., 2004[3]	G1: Control (not abstracted) G2: Guideline (increase reach) G3: Guideline + reminder notice and stickers for patients' charts (multi-component)	Convenience Physician computer-generated random number	Quebec physicians that had to be a primary prescribing physician (responsible for more than half of all anti-anginal prescriptions) for at least one patient (over 65 years old), and still be prescribing cardiovascular mediations as of 30 December 1999.	NR	Total:3293 G2: 1087 G3: 1115	Total: 2326 G2: 766 G3: 793	Total: 2326 G2: 766 G3: 793	

Table E-2. Key question 2 sample characteristics, part 1 (continued)

Author, Year	Groups	Sampling Strategy, Unit of Randomization, Process of Randomization	Inclusion/ Exclusion Criteria	N Eligible	N Randomized	N Completers	N Analyzed	Other Notes
Becker et al., 2008[4]	G1: Mailed guideline (Increase clinician reach) G2: Guideline implementation (multi-component, clinicians only) G3: Guideline implementation and motivational counseling directed at patient (multi-component, clinicians and patients)	General Practice: consecutive Patient: consecutive - all patients with low back pain General Practice General Practice: central permuted block randomization with allocation concealment	General Practice Inclusion criteria for practices were the willingness to participate of at least one physician and one practice nurse. Patient: Inclusion criteria for patients were LBP as presenting symptom on the day of recruitment, written consent to participate in the study, and age above 19 years. Exclusion criteria were insufficient German language skills, pregnancy, and isolated thoracic pain.	Practice: N =118 Patient: N = 1588	Practice N = 116 Patient N = 1378	Practice N= 116 Patient Baseline N = 1378 G1: 479 G2: 489 G3: 410 6 months N = 1261 G1: 450 G2: 435 G3: 376 12 months N = 1211 G1: 425 G2: 421 G3: 365	Patient Baseline N= 1378 G1: 479 G2: 489 G3: 410 6 months N = 1261 G1: 450 G2: 435 G3: 376 12 months N = 1211 G1: 425 G2: 421 G3: 365	

Table E-2. Key question 2 sample characteristics, part 1 (continued)

Author, Year	Groups	Sampling Strategy, Unit of Randomization, Process of Randomization	Inclusion/ Exclusion Criteria	N Eligible	N Randomized	N Completers	N Analyzed	Other Notes
Bekkering et al., 2005[5, 40]	G1: Received guidelines by mail (increase reach) G2: Received guidelines + active training strategy (multi-component)	All practices, convenience physical therapists, consecutive patients Practice location Block randomization. Block randomization (blocks of four practices) was carried out after pre-stratification for the work setting (solo/duo practices versus group practices). A statistician, who was not involved in this trial, drew up an allocation schedule using a computerized random number generator. The primary investigator (GEB), without any knowledge of the practices, listed them alphabetically according to the name of their street address, and subsequently assigned them to the intervention or control group using the allocation schedule.	Physio-therapists were eligible for participation if they worked in a private practice in primary care and if they expected to treat at least five patients with low back pain during the enrolment period. Included physical therapy practices that were members of the Royal Dutch Society for Physical Therapy (KNGF) located in or around the cities of Utrecht, Amersfoort, and Hilversum in the Netherlands Patients were eligible for inclusion if the physiotherapist confirmed that the diagnosis was non-specific low back pain and if the patient was able to complete questionnaires in the Dutch language.	N = 325 practices	Practices N = 68 G1: 34 G2: 34 Physio-therapists G1: 61 G2: 52 Overall: 511 patients G1: 259 patients G2: 256 patients	Physio-therapists G1: 48 G2: 37 Patients G1: 253 G2: 247 Patients: Baseline Overall: 483* G1: 241 G2: 242 6 weeks Overall: 465* G1: 230 G2: 235 12 weeks Overall: 448* G1: 223 G2: 225 26 weeks Overall: 439* G1: 221 G2: 218 52 weeks Overall: 428* G1: 214 G2: 214	Physio-therapists G1: 48 G2: 37 Patients G1: 253 G2: 247 PRIMARY ANALYSES (INTENT TO TREAT) Patients: Baseline Overall: 511 patients G1: 259 patients G2: 256 patients 6 weeks Overall: 511 patients G1: 259 patients G2: 256 patients 12 weeks Overall: 511 patients G1: 259 patients G2: 256 patients 26 weeks	

Table E-2. Key question 2 sample characteristics, part 1 (continued)

Author, Year	Groups	Sampling Strategy, Unit of Randomization, Process of Randomization	Inclusion/ Exclusion Criteria	N Eligible	N Randomized	N Completers	N Analyzed	Other Notes
Bekkering et al., 2005[5, 40] (continued)			Patients who were pregnant were excluded, as were those considered by the physiotherapist to be at high risk for dropping out of the study due to psychological problems.				Overall: 511 patients G1: 259 patients G2: 256 patients 52 weeks Overall: 511 patients G1: 259 patients G2: 256 patients	
Bishop and Wing, 2006[6]	G1: Control (not abstracted) G2: Physician only (increase reach) G3: Physician and patient (multi-component)	Unclear. Possibly consecutive (b/c used random number generator). Physician + patient Random number generator	Inclusion: Residents of British Columbia, Canada, aged between 19 and 65 years. They had as their chief complaint, acute low back pain and an accepted claim with the Workers' Compensation Board of British Columbia relating to an injury that was thought to be causative. All patients included in the study satisfied the Quebec Task Force Classification of Spinal Disorders criteria for categories 1 or 2 and had symptoms for more than 2 weeks and less than 4 weeks.	NR	Overall N: 462	Overall N: 428 G2: 149 G3: 139	0-4 weeks Overall: 462 G2: 162 G3: 151 5-12 weeks Overall N: 448 G2: 154 G3: 145 >12 weeks Overall N: 428 G2: 149 G3: 139	

Table E-2. Key question 2 sample characteristics, part 1 (continued)

Author, Year	Groups	Sampling Strategy, Unit of Randomization, Process of Randomization	Inclusion/ Exclusion Criteria	N Eligible	N Randomized	N Completers	N Analyzed	Other Notes
Bishop and Wing, 2006[6] (continued)			Exclusion: Demonstrated clinical findings on physical examination or had diagnostic imaging findings that suggested any other significant spinal pathology or co-morbidity that might independently influence the primary outcome. Patients were also excluded if they had persisting pain in any other areas of their spine, signs of systemic infection, malignancy, or pregnancy.					
Campbell et al., 2004[7]	G1: Control (not abstracted) G2: LHA (increase motivation) G3: TPV (multi-component) G4: TPV and LHA (multi-component)	Random Church Independent statistician randomly assigned into 4 groups	Churches had to have 80 or more active members Participants had to be active members over 18 years old	N = 26 N = 1463	N = 12 N = 850	N = 587 G2: 123 G3: 159 G4: 176	N = 587 G2: 123 G3: 159 G4: 176	

Table E-2. Key question 2 sample characteristics, part 1 (continued)

Author, Year	Groups	Sampling Strategy, Unit of Randomization, Process of Randomization	Inclusion/ Exclusion Criteria	N Eligible	N Randomized	N Completers	N Analyzed	Other Notes
Carney et al., 2005[8]	G1: Mailed health information (increase reach) G2: Telephone counseling (increase motivation)	Consecutive Patients NR	Eligible participants were consenting NH women aged 50 and older whose first mammogram recorded in the registry occurred between 1 May 1996 and 30 April 1996. Women with a personal history of breast cancer and women whose initial screening mammogram was reported as abnormal were excluded because we chose to focus the study on a population eligible for routine screening	Overall N: 300 (42 women could not complete all aspects of the study)	Overall N: 258 G1: 126 G2: 132	Overall N: 258 G1: 126 G2: 132	Overall N: 258 G1: 126 G2: 132	
Christakis et al., 2006[9]	G1: Usual care (not abstracted) G2: Parental content Alone (increase reach) G3: Provider notification alone (not abstracted) G4: Parental content and provider notification (multi-component)	Convenience Child (patient) NR	Eligible children were <11 years of age, had parents who spoke English, were patients at a participating clinic, and needed to make a well-child visit during the study period.	Overall N: 2209	Overall N: 887 G3: 211 G4: 210	Overall N: 767 G3: 183 G4: 177	Overall N: 767 G3: 183 G4: 177	

Table E-2. Key question 2 sample characteristics, part 1 (continued)

Author, Year	Groups	Sampling Strategy, Unit of Randomization, Process of Randomization	Inclusion/ Exclusion Criteria	N Eligible	N Randomized	N Completers	N Analyzed	Other Notes
Davis et al., 2004[10]	G1: Control - guidelines by mail (increase reach) G2: Intermediate (multi-component) G3: High intervention (multi-component)	Convenience Practice location Computer-generated random numbers were used, by a researcher not connected with the trial, to randomize locations (clusters) to control, intermediate, and intensive intervention arms.	Practice: All general practices in Tayside were eligible to participate (except four practices that had participated in earlier pilot projects). Patient-level: Patients on the lists of participating GPs who were receiving medication for epilepsy and were older than 16 years were eligible to take part.	Practice: Overall: 71 Patient Overall: 2, 025	Practice: Overall: 68 practices in 53 locations G1: 18 locations, 24 practices G2: 18 locations, 22 practices G3: 17 locations, 22 practices	Patients: Overall N: 811 G1: 255 G2: 269 G3: 287	Patients at Baseline: Overall:1,133 G1: 370 G2: 364 G3: 399 Patients at follow-up Overall: 811 G1: 255 G2: 269 G3: 287	
Eaton et al., 2011[11]	G1: 1-hour academic detailing (increase clinician ability) G2: Academic detailing plus a patient education toolkit, a computer kiosk with patient activation software, and a PDA-based decision support tool (multi-component)	Consecutive patients Primary care practices Practices were block randomized by size, specialty, and percentage of patients at LDL goal	Primary care practices in Southeastern New England; other inclusion criteria not reported	Overall N: 79 G1: NR G2: NR	Overall N: 30 primary care practices G1: 15 G2: 15 55 primary care physicians G1: 29 G2: 26 Overall N: 4,239* G1: 2,161 G2: 2,078	Overall N: 30 primary care practices G1: 15 G2: 15 55 primary care physicians G1: 29 G2: 26 4,105 patients G1: 2,105 G2: 2,000	Overall N: 30 primary care practices G1: 15 G2: 15 55 primary care physicians G1: 29 G2: 26 4,105 patients G1: 2,105 G2: 2,000	

Table E-2. Key question 2 sample characteristics, part 1 (continued)

Author, Year	Groups	Sampling Strategy, Unit of Randomization, Process of Randomization	Inclusion/ Exclusion Criteria	N Eligible	N Randomized	N Completers	N Analyzed	Other Notes
Elder et al., 2005[12]	G1: Culturally targeted print-materials + activity inserts (increase reach) G2: Tailored print materials + activity inserts + supporting materials (multi-component). G3: Tailored print materials + in-person promotora (multi-component)	Random Individual Randomly assigned participants using block randomization	Inclusion: Spanish-language dominant women between 18 and 65 Exclusion: there was no adult female living in the home; there was no adult female between 18 and 65 years of age; or the target adult female was pregnant, on a special diet for medical reasons, or planning to leave the San Diego area during the study period.	Authors don't provide exact # of eligible, but they say that over 1/3 of the 2,572 recruited women were not eligible	N = 357 G1: 119 G2: 118 G3: 120	N = 313 G1: 107 G2: 99 G3: 107	N = 313 G1: 107 G2: 99 G3: 107	

Table E-2. Key question 2 sample characteristics, part 1 (continued)

Author, Year	Groups	Sampling Strategy, Unit of Randomization, Process of Randomization	Inclusion/ Exclusion Criteria	N Eligible	N Randomized	N Completers	N Analyzed	Other Notes
Feldstein et al., 2006[13]	G1: Usual care (not abstracted) G2: EMR reminder (increase reach for clinicians) G3: EMR reminder and patient reminder (via letter with educational materials (multi-component)	Random Patients Study statistician randomized and assigned participants to the study groups using a design-adaptive randomization that balanced on age and fracture type. Computer random-number generator seeded by date and time generated a random sequence	Inclusion: individuals aged 50-89 as of study start date, HMO members least 12 months before study start date, individuals with study-defined fracture in 1999. Exclusions: Individuals having received pharmacological treatment for osteoporosis, BMD measurement, or exclusionary medical condition (malignancies (except melanoma skin cancers), chronic renal failure, dementia, organ transplant and cirrhosis in 12 months before start of study. Other exclusions: men, those without primary care provider, participants in osteoporosis clinical trials, nursing home residents, those without an address, and research center employees.	Overall N: 327 G1: NR G2: NR	Overall N: 327 G1: 107 G2: 107 G3: 107	Overall N: 314 G1: 103 G2: 101 G3: 110	Overall N: 311 G1: 101 G2: 101 G3: 109	Age reported in 9 year increments, personal income in 3 categories, education in 4 categories, no race or ethnicity information, no insurance information although all patients were part of an HMO so we can likely assume all had some coverage

Table E-2. Key question 2 sample characteristics, part 1 (continued)

Author, Year	Groups	Sampling Strategy, Unit of Randomization, Process of Randomization	Inclusion/ Exclusion Criteria	N Eligible	N Randomized	N Completers	N Analyzed	Other Notes
Gattellari et al., 2005[14]	G1: Leaflet (increase reach) G2: Video (increase reach) G3: Booklet (increase reach)	Random sample Participant Block randomization via computer software; participants and interviewers were blinded to allocation, but it's not possible to blind the interviewers at the post-test because they had to check to see if the participants had received the information	Aged 50-70 years old males with no known history of prostate CA and fluent in English. If more than 1 eligible respondent per household, respondent chosen at random.	Overall N: 585 Groups NR*	Overall N: 421 G1: 140 G2: 141 G3: 140	Overall N: 405 G1: 136 G2: 138 G3: 131	Overall N: 405 G1: 136 G2: 138 G3: 131	They present their demo-graphics from the pre-test and included people they didn't analyze
Hagmolen et al., 2008[15]	G1: Guideline dissemination (increase reach) G2: Guideline dissemination + educational program (increase ability) G3: Guideline dissemination + educational program + individualized treatment advice based on airway responsiveness and symptoms (multi-component)	Convenience Health center NR	Children aged 7 to 17 years old, had at least two prescriptions of β2-agonists or ICS were prescribed in the year before invitation. Children who were also treated by a pediatrician or pulmonologist were excluded as were children who had a disability, other relevant diseases, conductive disorders, or disturbing psychological problems.	Cluster: 18 Participants: 539	Cluster: 18 Participants: 404 G1: 114 G2: 143 G3: 147	Cluster: 18 Participants: 362 G1: 98 G2: 133 G3: 131	Cluster: 18 Participants: 362 G1: 98 G2: 133 G3: 131	

Table E-2. Key question 2 sample characteristics, part 1 (continued)

Author, Year	Groups	Sampling Strategy, Unit of Randomization, Process of Randomization	Inclusion/ Exclusion Criteria	N Eligible	N Randomized	N Completers	N Analyzed	Other Notes
Jain et al., 2006[16]	G1: Passive intervention-guidelines by mail (increase reach) G2: Active intervention (multi-component)	Convenience ICU clusters (ICUs that shared staff and procedures with other ICUs were grouped into clusters and were stratified according to size and setting) Randomization was computer generated and blinding was not possible.	Practice: ICU with at least 8 beds and a dietician Patient: all mechanically ventilated ICU patients who remained in the ICU for > 72 hours	Practice: Overall: 79	Practice Overall: 58 ICUs randomized as 50 clusters G1: 25 clusters G2: 25 clusters	Practice Overall: 58 ICUs randomized as 50 clusters G1: 25 clusters G2: 25 clusters	Practice Overall: 58 ICUs randomized as 50 clusters G1: 25 clusters G2: 25 clusters Patients Baseline Overall: 623 G1: 298 G2: 325 Follow-up Overall: 612 G1: 305 G2: 307 Note: the patients were not the same at baseline and follow-up. The authors took a cross-sectional survey at both time points.	

Table E-2. Key question 2 sample characteristics, part 1 (continued)

Author, Year	Groups	Sampling Strategy, Unit of Randomization, Process of Randomization	Inclusion/ Exclusion Criteria	N Eligible	N Randomized	N Completers	N Analyzed	Other Notes
Jousimaa et al., 2002[17]	G1: Computerized version of guidelines (increase ability) G2: Textbook-based version of guidelines (increase reach)	Consecutive Physician Randomized centrally using computer-generated numbers	Included only recently qualified physicians who would work in a Finnish health center for at least 2 months during the study period from February 1998 until September 1999	Overall N: 209 physicians	Overall N: 139 physicians G1: 72 G2: 67	Physicians: Overall N: 130* G1: 66 G2: 64 Patient encounters: Overall N: 4,633 G1: 2,453 G2: 2,180	Patient encounters: Overall N: 3,484 G1: 1,793 G2: 1,691	
Junghans et al., 2007[18]	G1: Conventional guideline (increase reach) G2: Ratings about specific patients in vignettes (increase motivation)	Convenience Physician A research assistant randomized 363 physicians using minimization software to balance recruitment by the 9 centers and the 2 clinical specialties.	Members of the British Cardiac Society or were general practitioners in the primary care trusts referring to 9 cardiothoracic centers in England and Scotland.	N = 3238	N = 363 G1: 184 G2: 179	N = 292 G1: 147 G2: 145	N = 292 G1: 147 G2: 145	

Table E-2. Key question 2 sample characteristics, part 1 (continued)

Author, Year	Groups	Sampling Strategy, Unit of Randomization, Process of Randomization	Inclusion/ Exclusion Criteria	N Eligible	N Randomized	N Completers	N Analyzed	Other Notes
Kennedy et al., 2003[19]	G1: Control (not abstracted) G2: Information (increase reach) G3: Interview (increase motivation)	Convenience Patients Random allocation to one of the three groups was then carried out using a form of random permuted blocks, with block size randomly set to three, six or nine to avoid any possibility of selection bias. The allocation sequence was generated by computer and stratified by consultant and the age at which the woman left full-time education. Secure randomization was ensured by using a central telephone randomization system based at the study administration center.	Women consulting one of 28 consultant gynecologists from six hospitals in the south west of England. All women referred from primary to secondary care with uncomplicated menorrhagia, deemed non-urgent by their consultant, were considered for trial entry if their referral related to a new episode of menorrhagia.	N = 1301	N = 894 G2: 296 G3: 300	Overall N: 625 G2: 206 G3: 215	Baseline Overall N: 885 G2: 293 G3: 298 Post-consultation Overall: 717 G2: 244 G3: 236 Short-term follow-up Overall: 631 G2: 205 G3: 221 Long-term follow-up Overall N: 625 G2: 206 G3: 215	

Table E-2. Key question 2 sample characteristics, part 1 (continued)

Author, Year	Groups	Sampling Strategy, Unit of Randomization, Process of Randomization	Inclusion/ Exclusion Criteria	N Eligible	N Randomized	N Completers	N Analyzed	Other Notes
King et al., 2007[20]	G1: Attention control (not abstracted) G2: Counselor via phone (increase motivation) G3: Automated counselor via phone (increase reach)	Consecutive Individual Following stratification by gender, randomly allocated using computerized version of Efron procedure where	Eligibility criteria: 1) ages 55 years and older; 2) not initially engaged in more than 60 minutes per week of moderate-intensity or more vigorous physical activity over previous 6 months; 3) free of any medical condition that would limit participation in moderate-intensity exercise; 4) BMI 40; 5) average alcohol intake 3 drinks per day; 6) able to speak and understand English sufficiently to provide informed consent and participate in study intervention and assessment procedures; 7) regular access to a touchtone phone; 8) not planning to move from area over study period; and 9) willing to be randomized to any of 3 study arms.	N= 370	N = 218 G2: 73 G3: 75	N = 189 G2: 66 G3: 61	N = 189 G2: 66 G3: 61	

Table E-2. Key question 2 sample characteristics, part 1 (continued)

Author, Year	Groups	Sampling Strategy, Unit of Randomization, Process of Randomization	Inclusion/ Exclusion Criteria	N Eligible	N Randomized	N Completers	N Analyzed	Other Notes
Laprise et al., 2009[21]	G1: CME (increase ability) G2: CME + practice enablers and reinforcers (multi-component)	Convenience General Practitioner Computer generated list of random numbers. Random assignment was within strata defined by regions /nurses (n = 5) and estimated time available per patient encounter (< 18 min/pt or ≥ 18 min/pt).	All practicing GPs in 5 regions of Quebec. Had to see at least 25 patients ≥ 55 yrs in any 2 week period in a community setting.	GPs N = 142	N: 133 G1: 66 G2: 67 Patients Overall: 2344 G1: 948 G2: 1396	N: 131 G1: 66 G2: 65	N = 122 G1: 61 G2: 61	

Table E-2. Key question 2 sample characteristics, part 1 (continued)

Author, Year	Groups	Sampling Strategy, Unit of Randomization, Process of Randomization	Inclusion/ Exclusion Criteria	N Eligible	N Randomized	N Completers	N Analyzed	Other Notes
Lien et al., 2007,[22] Svetkey et al., 2003,[23] Young et al., 2009[24]	G1: Advice only (increase reach) G2: Advice + behavioral counseling using established intervention (multi-component) G3: Established intervention + DASH dietary recommendations (multi-component)	Convenience Patients Randomization assignments were made centrally by a computer program. Clinical staff then notified participants of their assigned group. Used random number, stratified by clinic and hypertension status. Blocks of 24	Population consisted of generally healthy adults with above optimal BP (120 to 139 mm Hg systolic and/or 80 to 89 mm Hg diastolic) and individuals with stage-1 hypertension (140 to 159 mm Hg systolic and/or 90 to 95 mm Hg diastolic) who met national criteria for a 6-month trial of nonpharmacological therapy. Persons were eligible if they fit above SBP and DBP criteria and were not taking antihypertensive medication. Other inclusion criteria were age >25 years and BMI 18.5 to 45.0 kg/m2. Major exclusion criteria were regular use of drugs affecting BP, history of target organ damage and/or diabetes, use of weight-loss medications, previous cardiovascular event, heart failure, and angina.	Overall N: 810 G1: 273 G2: 268 G3: 269	Overall N: 810 G1: 273 G2: 268 G3: 269	Overall: 765 G1:259 G2: 253 G3: 253 Different number of completers for each outcome, for individuals w/out BP at 6-month assessment and for those who had been taking BP meds, 3-month BP measurements were carried forward; if a 3-month BP measurement was unavailable, values were imputed using a "hot deck" procedure that drew values from participants in the advice-only group	Overall N: 810 G1: 273 G2: 268 G3: 269	

Table E-2. Key question 2 sample characteristics, part 1 (continued)

Author, Year	Groups	Sampling Strategy, Unit of Randomization, Process of Randomization	Inclusion/ Exclusion Criteria	N Eligible	N Randomized	N Completers	N Analyzed	Other Notes
Marcus et al., 2009[25]	G1: Contact control treatment delayed group (not abstracted) G2: Telephone-based individualized feedback (increase motivation) G3: Print-based individualized feedback (increase reach)	Random Individual NR	Inclusion: healthy, age 18-65, and underactive (i.e., participating in moderate or vigorous physical activity for 90 minutes or less per week. Exclusion: BMI of greater than 35, asthma, emphysema, chronic bronchitis, hypertension, heart disease, abnormal electrocardiogram, stroke, prescription medication that might impair exercise performance, chronic infectious disease, significant musculoskeletal problems or any other serious medical condition that might make exercise unsafe, pregnancy or plans to attempt pregnancy, self-report of more than three alcoholic drinks per day on 5 or more days per week, hospitalization for a psychiatric disorder in the last 6 months, or currently suicidal, bipolar, or psychotic.	Overall N: 1700 G1: NR G2: NR G3: NR	Overall N: 239 G1 (control - labeled wellness in table): 78 G2 (phone): 80 G3 (print) : 81	Overall N: 218 at 6 months; 205 at 12 months G1: 72 at 6 months, 69 at 12 months G2: 75 at 6 months, 70 at 12 months G3: 71 at 6 months, 66 at 12 months	Overall N: 218 at 6 months; 205 at 12 months G1: 72 at 6 months, 69 at 12 months G2: 75 at 6 months, 70 at 12 months G3: 71 at 6 months, 66 at 12 months	

Table E-2. Key question 2 sample characteristics, part 1 (continued)

Author, Year	Groups	Sampling Strategy, Unit of Randomization, Process of Randomization	Inclusion/ Exclusion Criteria	N Eligible	N Randomized	N Completers	N Analyzed	Other Notes
Maxwell et al.,2010[26]	G1: Control (not abstracted) G2: Educational session + letter to provider (multi-component) G3: Educational session + letter to provider + FOBT kit (multi-component)	Convenience Related groups of individuals from study site NR	Inclusion: member of 1 of 45 Filipino community-based organizations or churches aged 50-70; not current with CRC screening Exclusion: history of CRC	Overall N: 614	Overall N: 548 G1: 163 G2: 183 G3: 202	Overall N: 432 G1: 130 G2: 146 G3: 156	Overall N: 548 G1: 163 G2: 183 G3: 202	Several of baseline character-istics approach significant difference, in particular group G2 seems different from G1 or G3
Murtaugh et al.,2005[27]	G1: Usual care (not abstracted) G2: Basic intervention e-mail reminder (increase reach) G3: Augmented intervention of e-mail reminder + package of supporting materials (multi-component)	All Nurse NR	Included patients with: primary diagnosis of CHF (ICD9-CM 428), age 18 or older, English or Spanish-speaking, able to provide informed consent; excluded a small number of nurses missing the practice measures since the records for their patients were not available at the time of chart review and a small number of nurses who had only an initial visit with that patient	Overall N: 388 nurses G1-G3: NR	Overall N: NR	Overall N: 354 G1: 122 G2: 114 G3: 118	Overall N: 354 G1: 122 G2: 114 G3: 118	Significant differences in the groups by percent female and educational level

Table E-2. Key question 2 sample characteristics, part 1 (continued)

Author, Year	Groups	Sampling Strategy, Unit of Randomization, Process of Randomization	Inclusion/ Exclusion Criteria	N Eligible	N Randomized	N Completers	N Analyzed	Other Notes
Paradis et al.,2011[28]	G1: Paper handouts (increase reach) G2: Educational DVD (increase reach)	Convenience Parents Used a random-numbers Table E-to generate the assignment of weeks to treatment or control in blocks of 8 weeks to account for seasonal variations in the birthrate; research staff recruiting subjects in the waiting room was unaware of the treatment assignment for each week	Parents or primary caregivers ≥ 18 years old of newborns ≤ 1 month old presenting for their first visit Other inclusion criteria: parent knowledge of written and spoken English and access to a phone	Overall N: 244	Overall N: 137 participants G1: 67 G2: 70	Overall N: 131 G1: 64 G2: 67	Overall N: 131 G1: 64 G2: 67	
Partin et al., 2004[29]	G1: Usual care (not abstracted) G2: Pamphlet (increase reach) G3: Video (increase reach)	Convenience Patient Computer generated algorithm. Participants stratified by age, PSA history, and facility before randomization.	Male veterans age 50 and older who had no CaP and scheduled primary care appointment at one of four VA facilities in the Midwest between April and June 2001.	N = 1152	N = 1152 G2:384 G3: 384	N = 893 G2: 295 G3: 308	N = 893 G2: 295 G3: 308	

Table E-2. Key question 2 sample characteristics, part 1 (continued)

Author, Year	Groups	Sampling Strategy, Unit of Randomization, Process of Randomization	Inclusion/ Exclusion Criteria	N Eligible	N Randomized	N Completers	N Analyzed	Other Notes
Rahme et al., 2005[30]	G1: No treatment control (not abstracted) G2: Decision tree (increase ability) G3: Workshop (increase ability) G4: Workshop + decision tree (multi-component)	Convenience Town Town was "randomly allocated"	All general practitioners registered in the 8 towns; all patients 65 years or older who filled a prescription for an NSAID, COX-2 inhibitor, or acetaminophen between May 2000 and June 2001 written by one of these general practitioners. Patients with osteoarthritis were those who had at least one diagnosis for osteoarthritis (ICD-9 code 715) in the previous 1215 days	NR	Physicians: 249 G2: 54 G3: 29 G4: 84	Physicians: 249 G2: 54 G3: 29 G4: 84	Patients—pre-intervention N = 3280 G2: 948 G3: 379 G4: 1048 Post-intervention G2: 831 G3: 317 G4: 969	

Table E-2. Key question 2 sample characteristics, part 1 (continued)

Author, Year	Groups	Sampling Strategy, Unit of Randomization, Process of Randomization	Inclusion/ Exclusion Criteria	N Eligible	N Randomized	N Completers	N Analyzed	Other Notes
Rebbeck et al., 2006[31]	G1: Dissemination of guidelines by mail (increase reach) G2: Implementation group (multi-component)	At clinic level: convenience/quota sample (based on the highest and lowest median cost per whiplash claim by the insurer) Patient level: convenience Physiotherapists (clinicians) Physiotherapists were stratified into low and high cost providers and the physiotherapists in each stratum were randomized into the dissemination or implementation group by an insurer. Interventions were coded so that the purpose of allocation was concealed from the insurer. Stratification was concealed from the trial center.	Criteria for physiotherapy clinics: 100 Clinics in 2 states in Australia that had seen at least 5 whiplash cases in the previous year; selected 24 of the highest and 24 of the lowest cost Criteria for patients: Those who presented to the clinic with acute whiplash; being over 18 years old; having been involved in a motor vehicle accident within the previous six weeks; having sustained whiplash-associated disorder Grade I-III; and giving informed consent	Patients: Overall N: NR G1: NR G2: NR Physio-therapists: Overall N: 48	Physio-therapists: Overall N: 27 G1: 13 G2: 14 Patients: Overall N: 99 G1: 28 G2: 71	Physio-therapists: Overall N: 18* G1: 5 G2: 13 Patients: Overall N: 93* G1: 26 G2: 67	Patients: Overall N: 93* G1: 26 G2: 67 Physio-therapists: G1:5 G2: 14 Baseline G1:12, G2: 14	

Table E-2. Key question 2 sample characteristics, part 1 (continued)

Author, Year	Groups	Sampling Strategy, Unit of Randomization, Process of Randomization	Inclusion/ Exclusion Criteria	N Eligible	N Randomized	N Completers	N Analyzed	Other Notes
Rebbeck et al., 2006[31] (continued)		Physiotherapists were blinded to the study hypothesis by being informed that they were randomized into one of two implementation groups.						
Rimer et al., 2001[32]	G1: No treatment control/ usual care (not abstracted) G2: Tailored print (increase reach) G3: Tailored print + telephone counseling (multi-component)	random Note: oversampled women "non-adherent" to mammograms at baseline (2/3) Patient NR	Inclusion: age 40-44 (might be 41-46- because Table E-1 has this age range) or age 50-54 (might be 51-56 because Table E-1 has this age range) Exclusion: being out of specified age range; having had or currently having breast cancer; no longer BCBS member	Overall N: 2165 G1-G3 assignment NR	Overall N: 1127 G1-G3 assignment NR	Overall N: 1127 G1: 412 G2: 392 G3: 323	Overall N: 1127 G1: 412 G2: 392 G3: 323	

Table E-2. Key question 2 sample characteristics, part 1 (continued)

Author, Year	Groups	Sampling Strategy, Unit of Randomization, Process of Randomization	Inclusion/ Exclusion Criteria	N Eligible	N Randomized	N Completers	N Analyzed	Other Notes
Simon et al., 2005[33]	G1: Mailed educational materials (increase reach) G2: Individual academic detailing (increase ability) G3: Group academic detailing (increase ability)	Clinics: Convenience; Clinicians and patients: Consecutive Clinics NR	All patients with hypertension receiving primary care at one of the 9 study sites; all clinicians providing primary care for adults at the 9 study sites	Clinics: Overall N: 9 G1: 3 G2: 3 G3: 3 Clinicians: Overall N: 781* G1: 319 G2: 235 G3: 227 Patients: Overall N: 9820 G1: NR G2: NR G3: NR	Clinics: Overall N: 9 G1: 3 G2: 3 G3: 3	Clinicians: Overall N: 781* G1: 319 G2: 235 G3: 227 Patients: Overall N: 9820 G1: NR G2: NR G3: NR	Clinicians: Overall N: 367 G1: 133 G2: 114 G3: 120 Patients: Overall N: 9820 G1: NR G2: NR G3: NR	

Table E-2. Key question 2 sample characteristics, part 1 (continued)

Author, Year	Groups	Sampling Strategy, Unit of Randomization, Process of Randomization	Inclusion/ Exclusion Criteria	N Eligible	N Randomized	N Completers	N Analyzed	Other Notes
Soler et al., 2010[34]	G1: Control (not abstracted) G2: Training session on the SEPAR guidelines (increase ability) G3: G2 + portable-device for spirometry (multi-component)	NR Clinicians NR	GPs: Practice at 1 of 40 included practices Patients: > 35 y/o, COPD or suspected COPD. First 5 patients/ GP recruited.	Overall N: 3254 physicians who selected 16,024 patients	Overall N: 2624 physicians (630 Declined to Participate) G1: 301 physicians selected 1481 patients G2: 1182 physicians selected 5798 patients G3: 1141 physicians selected 5556 patients	Overall N: 2624 physicians G1: 301 physicians selected 1481 patients G2: 1182 physicians selected 5798 patients G3: 1141 physicians selected 5556 patients	Overall N: 2624 physicians G1: 301 physicians selected 1481 patients G2: 1182 physicians selected 5798 patients G3: 1141 physicians selected 5556 patients	Patient characteristics presented were primary patient diagnosis (COPD, bronchial asthma, suspected COPD and others) and services received (forced spirometry, blood gases, chest x-rays)
Sullivan et al., 2010[35]	G1: VA guidelines (increase reach) G2: COPE: web-based education program (increase ability)	Consecutive Clinicians Residents were randomized in blocks according to gender and residency year to either COPE or the VA guidelines	Included residents in internal medicine; no reported exclusion criteria	Overall N: 570 residents	Overall N: 213 residents G1: 104 G2: 109	Overall N: 173* G1: 85 G2: 88	Overall N: 173* G1: 85 G2: 88	

Table E-2. Key question 2 sample characteristics, part 1 (continued)

Author, Year	Groups	Sampling Strategy, Unit of Randomization, Process of Randomization	Inclusion/ Exclusion Criteria	N Eligible	N Randomized	N Completers	N Analyzed	Other Notes
Watson et al., 2002[36]	G1: Guideline materials by postal mail (increase reach) G2: EO session and guidelines (increase ability) G3: CPE session and guidelines (increase ability) G4: Guidelines + EO and CPE (multi-component)	Random Pharmacies Random number assignment by statistician independent of the research team	NR	Overall N: 121 pharmacies	Overall N: 61* G1: 15 G2: 15 G3: 15 G4: 15 *Two pharmacies shared the same pharmacist and were therefore, randomized and treated as one pharmacy for the purpose of the study.	Overall N: 61* G1: 15 G2: 15 G3: 15 G4: 15	Overall N: 61* G1: 15 G2: 15 G3: 15 G4: 15	
Wetter et al., 2006[37]	G1: Single standard telephone-counseling session (increase reach) G2: Multiple enhanced telephone counseling sessions (multi-component)	Consecutive Smokers/individuals NR	Called the NCI's South Central CIS office to request smoking cessation help in Spanish; currently living in Texas; at least 18 years old; self-identification as a current smoker, Spanish speaking	Overall N: 355*	Overall N: 297 G1: 148 G2: 149	Overall N: NR G1: NR G2: NR	Overall N: NR G1: NR G2: NR	

Table E-2. Key question 2 sample characteristics, part 1 (continued)

Author, Year	Groups	Sampling Strategy, Unit of Randomization, Process of Randomization	Inclusion/ Exclusion Criteria	N Eligible	N Randomized	N Completers	N Analyzed	Other Notes
Wolters et al., 2005[38]	G1: Control mailed guidelines (increase reach) G2: Intervention involving package for learning, supporting materials, decision tree, and information leaflets for patients (multi-component)	Convenience Physician An independent statistician delivered computer-generated random numbers	None listed for GPs Patients All LUTS patients of older than 50 years visiting the GP. Exclusion criteria were: terminal phase of a disease, cognitive problems, known prostate carcinoma, a ureterostomy or bladder catheterization.	GP N = 142	GP N =142 G1: 72 G2: 70	GP N = 63 G1: 31 G2: 32 Patients N = 151 G1: 92 G2: 95	GP N = 63 G1: 31 G2: 32 Patients N = 151 G1: 92 G2: 95	
Wright et al., 2008[39]	G1: Standardized lecture by expert opinion leader (increase motivation) G2: Standardized lecture by expert opinion leader + academic detailing and a toolkit (multi-component)	Consecutive Hospitals Used a computer-generated scheme	Hospitals in Ontario included in study if they identified a local opinion leader in colon cancer	Overall N: 99 hospitals	Overall N: 42 hospitals G1: 21 G2: 21	Overall N: 34 hospitals G1: 18 G2: 16	Overall N: 34 hospitals G1: 18 G2: 16 Patients Overall N: 616 G1: 338 G2: 278	

Abbreviations: AF = audit and feedback; b/c = because; BCBS = Blue Cross Blue Shield; BMD = Bone Mineral Density; BMI = body mass index; BP = blood pressure; CA = cancer; CAL = computer-assisted learning; CHF = congestive heart failure; CIS = computer information systems; CME = continuing medical education; COPD = chronic obstructive pulmonary disease; COPE = Compassionate Options for Progressive Eldercare; COX-2 = Cyclooxygenase-2; CPE = continuing professional education; CRC = colorectal cancer; DASH = Dietary Approaches to Stop Hypertension; DBP = diastolic blood pressure; DVD = optical disc storage format; EMR = electronic medical record; EO = Education Outreach; FOBT = fecal occult blood test; G = group; GEB = G.E. Bekkering; GPs = general practitioners; HMO = health maintenance organization; ICD = International Classification of Diseases; ICD9-CM = International Classification of Diseases Ninth Revision Clinical Modification; ICU = intensive care unit; KNGF = Royal Dutch Society for Physical Therapy; LBP = lower back pain; LDL = low-density lipoprotein; LHA = lay health advisor; LUTS = lower urinary tract symptoms; min/pt = minute per patient; N = number; NCI = National Cancer Institute; NH = New Hampshire; NR = not reported; NSAID = nonsteroidal anti-inflammatory drug; PDA = personal device assistant; PSA = prostate-specific antigen; SBP = systolic blood pressure; SEPAR = Spanish Society of Pulmonology; TPV = tailored and targeted print and video; VA = Veterans Administration; Yrs = years

Table E-3. Key question 2 sample characteristics, part 2

Author, Year	Groups	Mean Age	Female	Race, Ethnicity	Median Income	Insured	Education	Health literacy/ numeracy	Others baseline characteristics (^=significant)
Bahrami et al., 2004[1]	G1: Mailed guideline (increase reach) G2: Guideline + AF (not abstracted) G3: CAL (increase ability) G4: CAL + AF (not abstracted)	42	10 24%	NR NR	NR	NR	NR	NR	Years of education
Banait et al., 2003[2]	G1: Mailed guidelines (increase reach) G2: Educational outreach (Multi-component)	NR	NR	NR NR	NR	NR	NR	NR	# of GPs Total population served Partnership size Fundholding status Dyspepsia drug usage Previous involvement in local guideline initiatives
Beaulieu et al., 2004[3]	G1: Control (not abstracted) G2: Guideline (increase reach) G3: Guideline + reminder notice and stickers for patients' charts (multi-component)	NR	0.295	NR NR	NR	NR	NR	NR	Medical training Years of experience Patient Medication # of visits to dr^ location of residence^

Table E-3. Key question 2 sample characteristics, part 2 (continued)

Author, Year	Groups	Mean Age	Female	Race, Ethnicity	Median Income	Insured	Education	Health literacy/ numeracy	Others baseline characteristics (^=significant)
Becker et al., 2008[4]	G1: Mailed guideline (Increase clinician reach) G2: Guideline implementation (multi-component, clinicians only) G3: Guideline implementation and motivational counseling directed at patient (multi-component, clinicians and patients)	M = 49.1*	n = 801; 58.1%*	NR NR	NR	NR	Overall N = 1207 13/12 yr = 186 (15%) 10 yr = 362 (30%) 9 yr = 506 (42%) Other graduation = 146 (12%) No qualification = 7 (.01%)	NR	Employment status Marital status
Bekkering et al., 2005[5]	G1: Received guidelines by mail (increase reach) G2: Received guidelines + active training strategy (multi-component)	Physio-therapist 40.7 Patient 45.3*	Physio-therapist 46 (43%) Patient 259 (51.8%*)	NR NR	NR	325.7 (65.1%)*^	NR	NR	Physiotherapist: Solo/duo practice^ Yrs experience Post-grad education on low back pain Patient # with paid job QBPDS score baseline pain intensity NRS score Baseline sick leave due to back pain

Table E-3. Key question 2 sample characteristics, part 2 (continued)

Author, Year	Groups	Mean Age	Female	Race, Ethnicity	Median Income	Insured	Education	Health literacy/ numeracy	Others baseline characteristics (^=significant)
Bishop and Wing, 2006[6]	G1: Control (not abstracted) G2: Physician only (increase reach) G3: Physician and patient (multi-component)	M = 38.2* Re-calculated age based on Table E-4a	133, 30.3%*	NR NR	NR	NR	NR	NR	Average age at injury
Campbell et al., 2004[7]	G1: Control (not abstracted) G2: LHA (increase motivation) G3: TPV (multi-component) G4: TPV and LHA (multi-component)	52	0.74	99% African American NR	NR	89% insured	25% some education beyond high school	NR	Marriage status BMI
Carney et al., 2005[8]	G1: Mailed health information (increase reach) G2: Telephone counseling (increase motivation)	40-79	100%	NR NR	NR	Insured: 232 (90%*)	< HS: 15 HS graduate: 81, 31.4%* Some college: 80, 31%* College grad: 73, 28%*	NR	Menopausal status HRT use Breast density Family history Marital status
Christakis et al., 2006[9]	G1: Usual care (not abstracted) G2: Parental content Alone (increase reach) G3: Provider notification alone (not abstracted) G4: Parental content and provider notification (multi-component)	4.1 years old *	46.5%*	White: 61.5%* Black: 7.3%* Asian: 6.5% Hispanic: 11%* Mixed: 11.8%* Other: 2.3 %* NR	(or parent) <10K: 6%* 10-25K: 14.3%* 25-50K: 27%* 50-75K: 26.5%* >75K: 26.5%*	NR	(Of Parent) < HS graduate:6.8%* HS graduate:17.5%* Some college:24.3%* College:33.8%* > college: 15.3%*	NR	Home internet access

Table E-3. Key question 2 sample characteristics, part 2 (continued)

Author, Year	Groups	Mean Age	Female	Race, Ethnicity	Median Income	Insured	Education	Health literacy/ numeracy	Others baseline characteristics (^=significant)
Davis et al., 2004[10]	G1: Control - guidelines by mail (increase reach) G2: Intermediate (multi-component) G3: High intervention (multi-component)	49.3*	53%*	NR NR	NR	NR	NR	NR	1 or more seizures per month
Eaton et al., 2011[11]	G1: 1-hour academic detailing (increase clinician ability) G2: Academic detailing plus a patient education toolkit, a computer kiosk with patient activation software, and a PDA-based decision support tool (multi-component)	Physicians: 46.6* Patients: 53.2*	Physi-cians: 17*, 30.9*% Patients: 2431.1* , 59.2*%	Patients: American Indian: 20.5 (0.5%) Asian: 30.8 (0.75%*) African American: 49.3 (1.2%*) White: 3928.5 (95.7%*) Patients: Hispanic: 61.6* , 1.5%*	NR	NR	NR	NR	Physicians: years in practice, patients seen per week, never used PDA, minutes behind at the end of the day Patients: martial status, CHD risk group, current smoker, physically inactive, at LDL goal, at non-HDL goal, diagnosed lipid disorder treated, treatment gap for lipip management, medical history, total cholesterol, LDL cholesterol, HDL cholesterol, prescription drugs, chronic conditions Practices: size, or nurse practitioners/ PAs in practice, type, hospital affiliated

Table E-3. Key question 2 sample characteristics, part 2 (continued)

Author, Year	Groups	Mean Age	Female	Race, Ethnicity	Median Income	Insured	Education	Health literacy/ numeracy	Others baseline characteristics (^=significant)
Elder et al., 2005[12]	G1: Culturally targeted print-materials + activity inserts (increase reach) G2: Tailored print materials + activity inserts + supporting materials (multi-component). G3: Tailored print materials + in-person promotora (multi-component)	39.71	1	NR 100% Latina	Median NR Income ($ per month) ≤ 1,000 = 13.3% 1,001-2,000 = 42.3% 2,001-3,000 = 29.9% > 3,000 = 14.5%	NR	0-6 years = 95, 26.6% Middle school =89, 24.9% High school = 76, 21.3% Some college = 97, 27.2%	NR	Country of formal education Employment status Self-perceived health^ Marital status Range of dietary variables Total family size Age BMI Waist-hip ratio
Feldstein et al., 2006[13]	G1: Usual care (not abstracted) G2: EMR reminder (increase reach for clinicians) G3: EMR reminder and patient reminder (via letter with educational materials (multi-component)	NR 50-89 age range Age reported in 9 year increments	100%	NR NR	NR personal income reported in 3 categories	NR No insurance information although all patients part of an HMO so can likely assume all had some coverage	G1: HS or less 31.7%; some college or more 22.8%; unknown 45.5% G2: HS or less 30.7%; some college or more 24.8%; unknown 44.6% G3:HS or less 35.8%; some college or more 25.7%; unknown 38.5% Education reported in 4 categories	NR	Fracture type, current smoker, weight, osteoporosis, Charlson Mobility Index, adequate total calcium (>1500 mg/d), regular activity

Table E-3. Key question 2 sample characteristics, part 2 (continued)

Author, Year	Groups	Mean Age	Female	Race, Ethnicity	Median Income	Insured	Education	Health literacy/ numeracy	Others baseline characteristics (^=significant)
Gattellari et al., 2005[14]	G1: Leaflet (increase reach) G2: Video (increase reach) G3: Booklet (increase reach)	58.1, range: 57.6-58.7 Presented their demographics from the pre-test and included people they didn't analyze	0	NR NR	NR	NR	(n=421) <high school grad: 15% high school degree or equivalent: 47% university degree: 37.5% Presented their demographics from the pre-test and included people they didn't analyze	NR	Employment status, occupation skill level; marital status*; Language usually spoken at home; self-reported health, PSA decision making preference; views towards PSA screening;
Hagmolen et al., 2008[15]	G1: Guideline dissemination (increase reach) G2: Guideline dissemination + educational program (increase ability) G3: Guideline dissemination + educational program + individualized treatment advice based on airway responsiveness and symptoms (multi-component)	10.8 *	160, 44%	NR NR	NR	NR	NR	NR	Duration of asthma Age at onset of asthma Lung function Asthma symptoms Asthma medication^ Atopic symptoms Asthma in relatives

Table E-3. Key question 2 sample characteristics, part 2 (continued)

Author, Year	Groups	Mean Age	Female	Race, Ethnicity	Median Income	Insured	Education	Health literacy/ numeracy	Others baseline characteristics (^=significant)
Jain et al., 2006[16]	G1: Passive intervention- guidelines by mail (increase reach) G2: Active intervention (multi-component)	61.5*	378, 61%*	NR NR	NR	NR	NR	NR	BMI Admission category Site characteristics
Jousimaa et al., 2002[17]	G1: Computerized version of guidelines (increase ability) G2: Textbook-based version of guidelines (increase reach)	Physicians: 27.1*	99, 71.2*%	NR NR	NR	NA	NA	NR	University, experience in health center (in months), type of health center (urban, rural), previous experience with EBMG guidelines (textbook, computer), mean # of searches by gender
Junghans et al., 2007[18]	G1: Conventional guideline (increase reach) G2: Ratings about specific patients in vignettes (increase motivation)	NR	NR	NR NR	NR	NR	NR	NR	Type of specialty # of years since qualification # of partners # of patients with angina per month
Kennedy et al., 2003[19]	G1: Control (not abstracted) G2: Information (increase reach) G3: Interview (increase motivation)	40.3*	100%	NR NR	NR	NR	NR	NR	Age on leaving full-time education Female consultant seen Median year of qualification of consultant Duration of problem Previous treatment Ever had surgery Knowledge Severity of menorrhagia Treatment preferences

Table E-3. Key question 2 sample characteristics, part 2 (continued)

Author, Year	Groups	Mean Age	Female	Race, Ethnicity	Median Income	Insured	Education	Health literacy/ numeracy	Others baseline characteristics (^=significant)
King et al., 2007[20]	G1: Attention control (not abstracted) G2: Counselor via phone (increase motivation) G3: Automated counselor via phone (increase reach)	60.77*	69.3%*	White: 87.4%* NR	NR	NR	Mean years of education: 16.2*	NR	Marriage status Employment status Body mass index # of meds
Laprise et al., 2009[21]	G1: CME (increase ability) G2: CME + practice enablers and reinforcers (multi-component)	GPs NR Patients 69.0	GPs 33, 27%* Patients 1043, 44%*	NR NR	NR	NR	NR	NR	GPs Region Yrs in practice Type of practice Patient Cardiovascular disease
Lien et al., 2007,[22] Svetkey et al., 2003,[23] Young et al., 2009[24]	G1: Advice only (increase reach) G2: Advice + behavioral counseling using established intervention (multi-component) G3: Established intervention + DASH dietary recommendations (multi-component)	50.0 (8.9)	500, 61.7%*	African American: 279 (34.4%)* Non-Hispanic White: 511 (63.9%)* All Others: 20 (2.5%)* NR	<$30,000: 84, 10.4%* $30,000-$60,000: 256, 31.6%* >$60,000: 441, 54.4%* Unknown: 29, 3.6%*	NR	High school or less: 74, 9.1%* Some college: 476, 58.8%* Some graduate school: 260, 32.1%*	NR	BMI Weight classification Alcohol, mean drinks Sedentary Current cigarette smokers Dyslipidemia Blood pressure Hypertensive
Marcus et al., 2009[25]	G1: Contact control treatment delayed group (not abstracted) G2: Telephone-based individualized feedback (increase motivation) G3: Print-based individualized feedback (increase reach)	44.5 years	82%	90.3 white NR	60.8% with total household income above $50,000		78.8% college graduate or postgrad work	NR	Marital status, employment, cigarette use, BMI, physical activity (min/week)

Table E-3. Key question 2 sample characteristics, part 2 (continued)

Author, Year	Groups	Mean Age	Female	Race, Ethnicity	Median Income	Insured	Education	Health literacy/ numeracy	Others baseline characteristics (^=significant)
Maxwell et al.,2010[26]	G1: Control (not abstracted) G2: Educational session + letter to provider (multi-component) G3: Educational session + letter to provider + FOBT kit (multi-component)	59 ± 6.2	363, 66%	100% Filipino American- NR	annual income <$50,000: 343 (62)* % corrected; wrong number in manuscript	384, 70%	College or more: 375 (68)	NR	Married, Baseline interview in English^, health problem, family history of cancer, regular doctor, MD recommended screening, ever had screening, recommended FOBT, had FOBT, MD recommended colonoscopy, ever had colonoscopy Note: Several of baseline characteristics approach significant difference, in particular group G2 seems different from G1 or G3

Table E-3. Key question 2 sample characteristics, part 2 (continued)

Author, Year	Groups	Mean Age	Female	Race, Ethnicity	Median Income	Insured	Education	Health literacy/ numeracy	Others baseline characteristics (^=significant)
Murtaugh et al.,2005[27]	G1: Usual care (not abstracted) G2: Basic intervention e-mail reminder (increase reach) G3: Augmented intervention of e-mail reminder + package of supporting materials (multi-component)	43.6	328*, 92.7 Significant differen-ces in the groups by percent female	White, non-hispanic: 82*, 23.2* Black, non-hispanic: 214*, 60.4* Hispanic: 25*, 7.1 Other: 33*,9.2*	NR	NR	Diploma: 47*, 13.3* Associate: 87*, 24.6* Completed college: 184*, 52.0 Postgrad: 15*, 4.2 missing: 21*, 5.9* Significant differences in the groups by educational level	NR	Percent per diem, means years of employment, number of eligible patients*
Paradis et al.,2011[28]	G1: Paper handouts (increase reach) G2: Educational DVD (increase reach)	36 (26%) < 21 years old 101 (74%) ≥ 21 years old	Infants: 68 (49.6%)* NR for parents/ caregivers	Infants: Black: 64 (47%) White: 16 (12%) Mixed/other: 57 (42%) NR for parents/ caregivers Infants: Hispanic: 38 (28%)	NR	Public insurance: 112 (82%)	< High school: 55 (40%) ≥ High school: 81 (60%)	NR	Parents/caregivers: first child, subject is mother Infant: hospital of birth, age at enrollment (days), gestational age (weeks), exclusively breastfed

Table E-3. Key question 2 sample characteristics, part 2 (continued)

Author, Year	Groups	Mean Age	Female	Race, Ethnicity	Median Income	Insured	Education	Health literacy/ numeracy	Others baseline characteristics (^=significant)
Partin et al., 2004[29]	G1: Usual care (not abstracted) G2: Pamphlet (increase reach) G3: Video (increase reach)	68.4	0	Nonwhite: 5.0% NR	NR	NR	< High School: 22.2% High School: 37.6% > High School: 40.2%	NR	Marriage status Overall health Comorbid conditions Prostate-specific items Urinary symptoms scale Medications
Rahme et al., 2005[30]	G1: No treatment control (not abstracted) G2: Decision tree (increase ability) G3: Workshop (increase ability) G4: Workshop + decision tree (multi-component)	76*	2264, 69%*	NR NR	NR	NR	NR	NR	At high risk for gastrointestinal events (%)
Rebbeck et al., 2006[31]	G1: Dissemination of guidelines by mail (increase reach) G2: Implementation group (multi-component)	Overall: 35.6 G1: 36.1 G2: 35.5	Overall: 79, 80% G1: 25, 89% G2: 54, 76%	NR NR	NR	99 (100%)	NR	NR	Patients: Dependents ^, Grade of Whiplash (I, II, III), duration of symptoms, mental health Physiotherapists: median cost / patient / therapist, whiplash caseload, knowledge of guidelines, location of therapist
Rimer et al., 2001[32]	G1: No treatment control/usual care (not abstracted) G2: Tailored print (increase reach) G3: Tailored print + telephone counselling (multi-component)	NR Age 41-46: 50% age 51-56: 50%	1127, 100%	Caucasian: 958*,85% African-American: 169*,15% NR	NR	100%	High school or less: 270,* 24% some college: 294* 35% college or more: 462,* 41%	NR	Married, work for pay, current smoker, perceived health as excellent or good at baseline, prior mammogram use; all NS differences

Table E-3. Key question 2 sample characteristics, part 2 (continued)

Author, Year	Groups	Mean Age	Female	Race, Ethnicity	Median Income	Insured	Education	Health literacy/ numeracy	Others baseline characteristics (^=significant)
Simon et al., 2005[33]	G1: Mailed educational materials (increase reach) G2: Individual academic detailing (increase ability) G3: Group academic detailing (increase ability)	Patients at baseline (N = 3692) Overall: NR G1: < 45 years: 24.4% 45-54 years: 30.3% 55-64 years: 22.1% 65-74 years: 15.7% ≥75 years: 7.5% G2: < 45 years: 18.8 45-54 years: 29.3 55-64 years: 25.5 65-74 years: 18.2 ≥75 years: 8.3 G3: < 45 years: 20.3 45-54 years: 27.6 55-64 years: 26.5 65-74 years: 16.7 ≥75 years: 8.9	1803.3, 48.8%*	NR NR	G1: $38,906 G2: $50,364 G3: $40,888	100%	No. and % high school education: 3577.7* , 96.9%*	NR	Insurance type, continuous health plan enrollment, diabetes, chronic disease score, rates of antihypertensive medication use

Table E-3. Key question 2 sample characteristics, part 2 (continued)

Author, Year	Groups	Mean Age	Female	Race, Ethnicity	Median Income	Insured	Education	Health literacy/ numeracy	Others baseline characteristics (^=significant)
Soler et al., 2010[34]	G1: Control (not abstracted) G2: Training session on the SEPAR guidelines (increase ability) G3: G2 + portable-device for spirometry (multi-component)	NR	32.9% of physic-cians	NR NR	NR	NR	NR	NR	Years of working life, type of center (rural, peri-urban, urban) and allocated population (<1500, 1500-2000, 2000-2500, >=2500) Note: Patient characteristics presented: primary patient diagnosis (COPD, bronchial asthma, suspected COPD and others) and services received (forced spirometry, blood gases, chest x-rays)
Sullivan et al., 2010[35]	G1: VA guidelines (increase reach) G2: COPE: web-based education program (increase ability)	NR	96, 45.1%	NR NR	NR	NA	NA	NR	Year of residency; residency program
Watson et al., 2002[36]	G1: Guideline materials by postal mail (increase reach) G2: EO session and guidelines (increase ability) G3: CPE session and guidelines (increase ability) G4: Guidelines + EO and CPE (multi-component)	NR	NR	NR	NR	NR	NR	NR	Location of practice, type of pharmacy

Table E-3. Key question 2 sample characteristics, part 2 (continued)

Author, Year	Groups	Mean Age	Female	Race, Ethnicity	Median Income	Insured	Education	Health literacy/ numeracy	Others baseline characteristics (^=significant)
Wetter et al., 2006[37]	G1: Single standard telephone-counseling session (increase reach) G2: Multiple enhanced telephone counseling sessions (multi-component)	41.1	133.1, 44.8%*	NR 198.1, 66.7*% Mexican origin	164.8 , 55.5*% <20,000	68.9, 23.2*%	% NR 10.9* years of education (less than a high school education)	NR	Marital status, employment status, immigrant status, language spoken at home , tobacco-related variables
Wolters et al., 2005[38]	G1: Control mailed guidelines (increase reach) G2: Intervention involving package for learning, supporting materials, decision tree, and information leaflets for patients (multi-component)	Gp 47.4 Patient 66.3	GP 16, 25.4% Patient 0%`	NR NR	NR	`NR	Patient Lower: 38.5% Secondary: 31.1% Higher: 23.2% Unknown: 7.3%	NR	GPs Years working as GP GP trainer Solo practice Rural area >1 hospital to refer to Patient Symptoms
Wright et al., 2008[39]	G1: Standardized lecture by expert opinion leader (increase motivation) G2: Standardized lecture by expert opinion leader + academic detailing and a toolkit (multi-component)	63.1* = Mean age at surgery	239 , 38.8%*	NR NR	NR	NR	NR	NR	Mean tumor size, mean specimen length, resection type, tumor stage T4

Abbreviations: AF = audit and feedback; BMI = body mass index; CAL = computer-assisted learning; CHD = coronary heart disease; CME = continuing medical education; COPD = chronic obstructive pulmonary disease; COPE = Compassionate Options for Progressive Eldercare; CPE = continuing professional education; DASH = Dietary Approaches to Stop Hypertension; dr = doctor; DVD = optical disc storage format; EBMG = European Board of Medical Genetics; EMR = electronic medical record; EO = Education Outreach; FOBT = fecal occult blood test; GPs = general practitioners; HDL = high-density lipoprotein; HMO = health maintenance organization; HRT = Hormone Replacement Therapy; HS =high school; LDL = low-density lipoprotein; LHA = lay health advisor; MD = medical doctor; NR = not reported; NRS = numeric rating scale; PA = physician's assistant; PDA = personal device assistant; PSA = prostate-specific antigen; QBPDS = Quebec Back Pain Disability Scale; SEPAR = Spanish Society of Pulmonology; TPV = tailored and targeted print and video; VA = Veterans Administration.

Table E-4. Key question 1 intervention descriptions

Author, Year	Groups	Comparators	Clinical Focus of the Evidence, Source of Information, Theory-Based, Prior Testing	Intervention Format, Delivery Agent, Intensity	Evidence Presentation	Message of Intervention
Bahrami et al., 2004[1]	G1: Mailed guideline (increase reach) G2: Guideline + AF (not abstracted) G3: CAL (increase ability) G4: CAL + AF (not abstracted)	Postal mail/email A copy of the guideline direct from SIGN plus a double sided laminated sheet known as the 'Quick Reference Guide' which summarizes the findings in an accessible manner. Skills building CAL intervention strategy consisted of a laptop computer based support tool, with the potential to assist dental practitioners in deciding on the appropriate treatment of third molars.	Management of impacted and unerupted third molars; treatment SIGN No No	G1: paper G3: electronic-based G1: postal G2: computer NR	Unclear	G1: Guideline recommendations + quick reference guide G3: The software was based solely on the SIGN guideline

Table E-4. Key question 1 intervention descriptions (continued)

Author, Year	Groups	Comparators	Clinical Focus of the Evidence, Source of Information, Theory-Based, Prior Testing	Intervention Format, Delivery Agent, Intensity	Evidence Presentation	Message of Intervention
Banait et al., 2003[2]	G1: Mailed guidelines (increase reach) G2: Educational outreach (Multi-component)	Postal mail/email Copy of guideline posted to all GPs in July 1997, 3 months prior to intervention Multi-component Personal visits by a trained person to health care providers in their own setting. Interactive educational workshops. Seminars held over a period of 6 weeks. Seminars involved 4 to 8 GPs from 2 to 3 practices. Presentation of guideline and then Q&A; info about available services, summaries of local data, copy of the text used during the discussions, contact details, reinforcement visit after 3 months.	Dyspepsia management; treatment Guideline; British Society of Gastroenterology dyspepsia management guideline No No	G1: paper-based G2: in-person G1: postal G2: in-person (local hospital specialists) Over 6 weeks; 90 minute meetings	NR	Guideline recommendations

Table E-4. Key question 1 intervention descriptions (continued)

Author, Year	Groups	Comparators	Clinical Focus of the Evidence, Source of Information, Theory-Based, Prior Testing	Intervention Format, Delivery Agent, Intensity	Evidence Presentation	Message of Intervention
Beaulieu et al., 2004[3]	G1: Control (not abstracted) G2: Guideline (increase reach) G3: Guideline + reminder notice and stickers for patients' charts (multi-component)	Postal mail/email One-page summary of the guideline Additional resources One-page summary of the guideline, followed a month later by a reminder notice, which included stickers to post on patients' charts	Anti-angina therapy; treatment Guideline; College des Medecins du Quebec No No	Paper-based Postal G2: 1 session G2: 2 sessions, 1 month apart	NR	Guideline recommendations. The key recommendations in the summary were: (1) to write a prescription for acetylsalicylic acid for patients with sTable E-angina; (2) to control serum cholesterol, with a target value for low-density lipoprotein cholesterol < 2.6 mmol/l; and (3) to favor b-blockers as the first choice for anti-angina medication. Data on prescribing rates for the three targeted medication classes by physicians practicing in the same regions as the participating physicians were also included in the one-page summary.

Table E-4. Key question 1 intervention descriptions (continued)

Author, Year	Groups	Comparators	Clinical Focus of the Evidence, Source of Information, Theory-Based, Prior Testing	Intervention Format, Delivery Agent, Intensity	Evidence Presentation	Message of Intervention
Becker et al., 2008[4]	G1: Mailed guideline (Increase clinician reach) G2: Guideline implementation (multi-component, clinicians only) G3: Guideline implementation and motivational counseling directed at patient (multi-component, clinicians and patients)	Postal mail/email Guideline delivered via postal mail. Targeting physician. Interpersonal outreach 3 interactive seminars with academic detailing and additional resources. Targeting physician. Champions Same as Comparator 2, except this group provided motivational counseling targeting toward the patient. Practice nurses were asked to invite all identified patients for up to 3 counseling sessions (max 10–15 minutes each), the first session within 1 to 3 weeks after inclusion in the study.	G1,G2, G3: Management of acute and chronic LBP—treatment G3: (For patients) Physical fitness—Prevention DEGAM No No	G1: paper-based G2: in-person + paper-based G3: in-person + paper-based G1: postal mail G2: in-person (study nurses) G3: in-person (trained practice nurses) G1: 1 session G2: 3 sessions and then 2 more academic detailing sessions after 3 to 6 months. G3: same as G2 + up to 3 counseling sessions, 10-15 minutes max.	Unclear	G1: Informational. A detailed version and a pocket card for physicians, a prescription-like short form information and a more detailed flyer for patients to be handed out during and after consultation. G2: & G3: Session 1 talked about performance of the diagnostic triage and identification of red flags. Session 2 identification of yellow flags, and general principles on management of chronic pain. Session 3 was informing and advising patients. Information about relevant resources for pain patients provided. Plus 2 individual educational visit by study nurses ("academic detailing"). They presented the guideline and at the second session they talked about problems with implementation.

Table E-4. Key question 1 intervention descriptions (continued)

Author, Year	Groups	Comparators	Clinical Focus of the Evidence, Source of Information, Theory-Based, Prior Testing	Intervention Format, Delivery Agent, Intensity	Evidence Presentation	Message of Intervention
Becker et al., 2008[4]						G3: Motivate LBP patients for regular physical activity. Nurses were encouraged to use specifically designed brochures on motivational and behavior change and posters to communicate the key messages.
Bekkering et al., 2005[5]	G1: Received guidelines by mail (increase reach) G2: Received guidelines + active training strategy (multi-component)	Postal mail/email Participants received guidelines by mail together with four forms: a self-evaluation form to assess whether their current management was consistent with the recommendations contained in the clinical guidelines, two forms facilitating discussion with other physiotherapists and general practitioners respectively, and a copy of the Quebec Back Pain Disability Scale. A summary of the clinical guidelines was also provided. Physical therapists re instructed "to act as usual," to read these guidelines if they have read previous guidelines and not read the guidelines if they have not read any other guidelines.	Low back pain; treatment Guidelines; the KNGF No No	G1: paper-based G2: in-person G1: postal G2: in-person (primary investigator and one of two additional trainers with adequate clinical experience in the management of low back pain) 2 sessions 2.5 hours each Total time: 5 hrs 2 hrs prep time before each session	Unclear	Guideline recommendations; educational; overcoming barriers G2: Session 1 included a didactic overview of the diagnostic and treatment processes: overview of the evidence and consequences of the evidence for diagnostic and therapeutic management compared with their own current management; interactive Q&A; two examples of role playing with an actor—one on the diagnostic process and one on the treatment process. A 4 week interval in which the physiotherapists were expected to implement the guidelines in practice.

Table E-4. Key question 1 intervention descriptions (continued)

Author, Year	Groups	Comparators	Clinical Focus of the Evidence, Source of Information, Theory-Based, Prior Testing	Intervention Format, Delivery Agent, Intensity	Evidence Presentation	Message of Intervention
Bekkering et al., 2005[5] (continued)		Multi-component Multifaceted program consisting of education, discussion, role playing, feedback, and reminders. Received guidelines by mail and active training strategy, which consisted of 2 training sessions with groups of 8-12 physical therapists. The aim of the sessions was to improve knowledge and skills regarding evidence-based physical therapy for patients with low back pain. Content of sessions was based on expected barriers to implementation. Sessions were interactive and involved group discussion, role playing, feedback, and reminders.				Session 2 consisted of a discussion of experiences with implementing the guidelines in practice; feedback on current management; two reminders with respect to evidence based patient education

Table E-4. Key question 1 intervention descriptions (continued)

Author, Year	Groups	Comparators	Clinical Focus of the Evidence, Source of Information, Theory-Based, Prior Testing	Intervention Format, Delivery Agent, Intensity	Evidence Presentation	Message of Intervention
Bishop and Wing, 2006[6]	G1: Control (not abstracted) G2: Physician only (increase reach) G3: Physician and patient (multi-component)	Additional resources Patient's family physician received a copy of the clinical practice guidelines with a letter from a study physician regarding a specific named patient encouraging compliance with the guidelines. In addition each family physician received a ''guideline reminder letter'' at each of three separate stages of the patient's clinical course summarizing the different aspects of the guidelines that specifically applied to the 0–4-week, 5–12-week, and greater than 12-week post injury periods Multi-component Family physician received the Group 2 intervention and in addition, the individual patient received lay language versions of a pamphlet outlining the clinical practice guidelines and of the same clinical practice guidelines reminders sent at the same time intervals of the patient's clinical course	Acute phase of a lower back injury; treatment Guidelines reviewed from 13 countries including the US (AHCPR) which were "remarkably consistent" No No	Paper-based Postal (?) # sessions:4 (baseline, 0-4 week; 5-12 week; > 12 week) Note: Hard to know exactly how clinicians "received" the intervention - was the "letter" handed to them or mailed?	NR Patients received a "lay language" version of the guideline.	Guideline recommendations

Table E-4. Key question 1 intervention descriptions (continued)

Author, Year	Groups	Comparators	Clinical Focus of the Evidence, Source of Information, Theory-Based, Prior Testing	Intervention Format, Delivery Agent, Intensity	Evidence Presentation	Message of Intervention
Campbell et al., 2004[7]	G1: Control (not abstracted) G2: LHA (increase motivation) G3: TPV (multi-component) G4: TPV and LHA (multi-component)	Social networks Church members submitted names of people in the church for whom they turned to for help and advice. Individuals who were named by multiple church members were identified as potential LHAs. These people were then invited to an orientation and were invited to volunteer as an LHA. If they accepted they were then trained. Multi-component Video + newsletter Multi-component Video + newsletter + lay health advisor	Colorectal cancer; prevention Guideline; USPSTF Yes Yes	G2: Paper based + video G3: In-person G4: Paper-based + video + in-person G2: Postal G3: In-person (church members) G4: Postal + in-person Baseline, 2, 4, 6, 9 months.	Mostly qualitative. Text files, graphics and photos in message library for tailoring.	G2: Newsletters were tailored. Newsletter personalized with names and included tailors elements, including feedback on dietary intake, physical activity, screening, CRC risk factors, social support, barriers. Message elements also targeted to cultural, spiritual, and community factors. Videos were targeted. Purpose of the videos was to provide additional motivating messages and modeling and skills demonstration to enhance and complement the information in the newsletters. Videos included testimonial and modeling skills. G3: LHA intervention designed to provide education and promote social support for behavioral change. LHA expected to organize and conduct at least 3 church-wide activities focused on spreading information and enhancing support for

Table E-4. Key question 1 intervention descriptions (continued)

Author, Year	Groups	Comparators	Clinical Focus of the Evidence, Source of Information, Theory-Based, Prior Testing	Intervention Format, Delivery Agent, Intensity	Evidence Presentation	Message of Intervention
Campbell et al., 2004[7] (continued)						health eating, physical activity, and CRC screening. All interventions based on the stages of change trans-theoretical framework, the health belief model, and social support models.

Table E-4. Key question 1 intervention descriptions (continued)

Author, Year	Groups	Comparators	Clinical Focus of the Evidence, Source of Information, Theory-Based, Prior Testing	Intervention Format, Delivery Agent, Intensity	Evidence Presentation	Message of Intervention
Carney et al., 2005[8]	G1: Mailed health information (increase reach) G2: Telephone counseling (increase motivation)	Postal mail/email Receive general health information packets by mail received a variety of brochures. Multi-component Telephone intervention based on the Trans-theoretical Model and contained both educational and motivational counseling components. Telephone intervention contained four components. First, while talking with study participants, health educators independently noted the woman's barriers to mammography and then, secondly, they assessed the woman's stage of readiness to change. Third step in intervention involved recording the details of the counseling provided to form the basis of the second counseling call that took place one year later. Final aspect of the intervention was to select the appropriate code(s) on the worksheet indicating barriers to screening and each woman's stage of change.	Breast cancer and mammography screening G1: USPSTF mammography screening recommendations, and a brochure describing services provided by the NH State Department of Health and Human Services as part of the NHBCCP. Yes No	G1: print-based G2: telephone G1: postal G2: female health educators G2: # sessions: 2 length: 6 min total time: 12 min	Unclear	G1: Screening recommendations and toll free numbers to find more information. G2: Motivational messages, Tailored information

Table E-4. Key question 1 intervention descriptions (continued)

Author, Year	Groups	Comparators	Clinical Focus of the Evidence, Source of Information, Theory-Based, Prior Testing	Intervention Format, Delivery Agent, Intensity	Evidence Presentation	Message of Intervention
Christakis et al., 2006[9]	G1: Usual care (not abstracted) G2: Parental content Alone (increase reach) G3: Provider notification alone (not abstracted) G4: Parental content and provider notification (multi-component)	Social networks Patients Alone. Parents in the group that received content alone were able to select topics and to access relevant content. However, their providers received no information regarding their use of the Web site. Social networks Patients + Physicians. Parents in the group that received Web content and notification were able to select and to read about topics in which they were interested. Providers had access to the topics in which the parents were interested and the results of any screening questionnaires they completed.	13 prevention topics (1) USPSTF Guide to Clinical Preventive Services, (2) Bright Futures guidelines for health supervision, (3) Peer-reviewed systematic reviews of other, preventive care interventions, and (4) High-quality, randomized, controlled trials. No No	Web-based Delivered in clinic or via web at home #: 1 length: NR total time: NR	Example sentences, it seems like it is both quantitatively and graphically, but that is unclear.	Provide relevant information (based on age-specific recommendations) that will inform and motivate patient to generate conversation with physician.

Table E-4. Key question 1 intervention descriptions (continued)

Author, Year	Groups	Comparators	Clinical Focus of the Evidence, Source of Information, Theory-Based, Prior Testing	Intervention Format, Delivery Agent, Intensity	Evidence Presentation	Message of Intervention
Davis et al., 2004[10]	G1: Control - guidelines by mail (increase reach) G2: Intermediate (multi-component) G3: High intervention (multi-component)	Postal mail/email Mailed a copy of the guideline Multi-component The intermediate group received the guideline, invitations to an interactive workshop (e.g., skills training), and structured protocol documents designed to supplement the guideline (e.g., provision of support materials). Multi-component Guideline + workshop and protocols + clinical nurse specialist	Epilepsy; diagnosis and treatment National evidence-based guideline on the Diagnosis and Management of Epilepsy in Adults, published by SIGN No Yes	G1: paper-based G2: paper-based + in-person G1: postal G2: postal + two consultant neurologists NR	NR	G1: Guideline recommendation G2: workshop was designed to go over case studies and promote discussion. The protocols were designed for use at the first presentation of a new patient and for use by either a practice nurse or a GP at regular review consultations. G3: Role of nurse was to offer advice and training to practices in establishing epilepsy review programs, to promote the use of the guideline in epilepsy management, and to provide information on epilepsy for both practitioners and patients.

Table E-4. Key question 1 intervention descriptions (continued)

Author, Year	Groups	Comparators	Clinical Focus of the Evidence, Source of Information, Theory-Based, Prior Testing	Intervention Format, Delivery Agent, Intensity	Evidence Presentation	Message of Intervention
Eaton et al., 2011[11]	G1: 1-hour academic detailing (increase clinician ability) G2: Academic detailing plus a patient education toolkit, a computer kiosk with patient activation software, and a PDA-based decision support tool (multi-component)	Academic detailing A 1-hour academic detailing session during which ATP III cholesterol guidelines were discussed and abbreviated guideline pocket guides were given to each physician; also received a PDA but without the decision support tool and had minimal further contact to mimic usual care Multi-component 1-hour academic detailing session (same as comparator #1); also, a patient education toolkit (and companion Web site), a computer kiosk with patient activation software, and a PDA-based decision support tool for each physician, which included 4 booster academic detailing sessions	Cholesterol treatment Guidelines No No	G1: In-person G2: In-person, Web-based, electronic-based NR G1:Academic detailing #: 1 length: 1 hour total time: 1 hour G2:Academic detailing #: 1 length: 1 hour Plus 4 booster sessions booster length: NR		Guideline recommendations

Table E-4. Key question 1 intervention descriptions (continued)

Author, Year	Groups	Comparators	Clinical Focus of the Evidence, Source of Information, Theory-Based, Prior Testing	Intervention Format, Delivery Agent, Intensity	Evidence Presentation	Message of Intervention
Elder et al., 2005[12]	G1: Culturally targeted print-materials + activity inserts (increase reach) G2: Tailored print materials + activity inserts + supporting materials (multi-component). G3: Tailored print materials + in-person promotora (multi-component)	Increased reach Targeted newsletters and activities inserts Multi-component Tailored newsletter and activities insert Multi-component Lay health advisor "Promotoras" + tailored newsletter and activities inserts	Reduce dietary fat and increase fiber Unspecified; American Heart Association NIH; American Dietetic Association, and the American Cancer Society Yes Unclear	G1 & G2: Paper-based G3: Paper-based + in-person G1 & G2: Postal G3: Promotoras (characteristics: Spanish-language dominant; naturally empathetic, able to develop rapport and to be neutral and nonjudgmental; perceived as a role model in the community; and interested in helping women change lifestyle behaviors.) G1: one time mailing (probably) G2: 12 weekly mailings G3: 12 weekly mailings of print materials + 12 weekly home visit or telephone call	NR	G1: Targeted materials developed for a Latino population and were available in Spanish. Language-appropriate materials that contained information on food purchasing, food preparation, and food consumption were available from the American Heart Association, American Dietetic Association, and the American Cancer Society G2: Newsletters provided feedback on the assessment process, as well as an opportunity for personalized goal setting and for dealing with identified barriers. Degree of complexity of the activity in the insert varied by the participant's readiness to change (e.g., acquire information vs. self-monitor). Participants encouraged to complete the activity on the insert and return the self-addressed stamped card to be entered into a raffle and to receive additional

Table E-4. Key question 1 intervention descriptions (continued)

Author, Year	Groups	Comparators	Clinical Focus of the Evidence, Source of Information, Theory-Based, Prior Testing	Intervention Format, Delivery Agent, Intensity	Evidence Presentation	Message of Intervention
Elder et al., 2005[12] (continued)						chapters of the story (novela) in the newsletter. Also magnetic flower petals containing healthy lifestyle messages and eight recipes. G3: Using the skills acquired in the program, as well as their natural ability to provide support and encouragement and their social networking skills, the promotoras worked with individual participants to negotiate behavioral change goals. The promotoras relied primarily on the participant's weekly tailored newsletter to guide discussions and suggest opportunities for skill development.

Table E-4. Key question 1 intervention descriptions (continued)

Author, Year	Groups	Comparators	Clinical Focus of the Evidence, Source of Information, Theory-Based, Prior Testing	Intervention Format, Delivery Agent, Intensity	Evidence Presentation	Message of Intervention
Feldstein et al., 2006[13]	G1: Usual care (not abstracted) G2: EMR reminder (increase reach for clinicians) G3: EMR reminder and patient reminder (via letter with educational materials (multi-component)	Multi-component G1: Patient-specific EMR in-basket messages for their enrolled patients from the chairman of the osteoporosis quality-improvement committee. Letter-style message informed the provider of the patient's risk of osteoporosis based upon the patient's age and prior fracture and state the need for evaluation and treatment; also listed internal and external guideline resources that provided detailed recommendations regarding evaluation, calcium and vitamin D intake, lifestyle and medication. Provider also given the option of contacting the sender for more information. Multi-component G2: (EMR reminder to clinician) plus patient reminder: a single mailing of an advisory letter with educational materials (addressing menopause, osteoporosis, calcium and	Osteoporosis National Osteoporosis Foundation, European Foundation for Osteoporosis and Bone Disease, American Association of Clinical Endocrinologists an d American College of Rheumatology Task Force on Osteoporosis Guidelines. Type of evidence NR. No Unclear	G1: Usual care G2: Electronic-based (email linked to EMR) G3: Electronic-based - EMR (clinician reminder) + postal mail print letter + provider receipt of patient letter Email for provider, mail/postal for patients G2: Baseline and 3 months after the first baseline message for providers. G3: One letter with educational material	Unclear	Options for reducing risk for osteoporosis, need for evaluation and treatment, information on calcium and vitamin D intake, lifestyle and medication

Table E-4. Key question 1 intervention descriptions (continued)

Author, Year	Groups	Comparators	Clinical Focus of the Evidence, Source of Information, Theory-Based, Prior Testing	Intervention Format, Delivery Agent, Intensity	Evidence Presentation	Message of Intervention
Feldstein et al., 2006[13] (continued)		vitamin D, physical activity, home safety and fall prevention). Letter identified patient's risk, discussed clinical practice guideline recommendations, and requested that the patient discuss management options with her provider.				
Gattellari et al., 2005[14]	G1: Leaflet (increase reach) G2: Video (increase reach) G3: Booklet (increase reach)	Postal mail/email G1: received a leaflet in the mail called "Testing for prostate cancer", 856 words in length, with a flesch reading score of 14.6 years It provided brief info on types of PSA screening tests; false-positives. No information on harms. Electronic/digital media G2 received a video that met criteria for a decision-aid; "The choice is yours: testing for prostate cancer'; 20 minutes in duration; younger man with family history of prostate ca and an older man individually weighing up the pros and cons of PSA screening; natural history of cancer, test accuracy, and treatments are described	Prostate cancer screening Cancer Foundation of Western Australia Yes No	Two groups, G1 and G3 received paper-based information while G2 received Postal #: 1 length: 20 minutes v. 2407 words total time: NR NOTE: Treatment, including statistics on treatment-related complications; included a visual aid in the form of a flow-chart outlining the consequences of screening and a	Quantitative and graphical	Included recommendations, prostate cancer risk, and testing risk and benefit.

Table E-4. Key question 1 intervention descriptions (continued)

Author, Year	Groups	Comparators	Clinical Focus of the Evidence, Source of Information, Theory-Based, Prior Testing	Intervention Format, Delivery Agent, Intensity	Evidence Presentation	Message of Intervention
Gattellari et al., 2005[14] (continued)		Postal mail/email G3: 2407 word evidence-based booklet, entitled "Should I have a PSA test for prostate cancer: information for men who want to know more about screening tests for prostate cancer; Flesch-Kincaid reading age of 11.8 years; included stats on the life-time and age-specifics risks of developing and dying from prostate cancer, family history as risk factor, test accuracy, and potential benefits and harms from treatment, including statistics on treatment-related complications; included a visual aid in the form of a flow-chart outlining the consequences of screening and a values clarification form.		values clarification form.		

Table E-4. Key question 1 intervention descriptions (continued)

Author, Year	Groups	Comparators	Clinical Focus of the Evidence, Source of Information, Theory-Based, Prior Testing	Intervention Format, Delivery Agent, Intensity	Evidence Presentation	Message of Intervention
Hagmolen et al., 2008[15]	G1: Guideline dissemination (increase reach) G2: Guideline dissemination + educational program (increase ability) G3: Guideline dissemination + educational program + individualized treatment advice based on airway responsiveness and symptoms (multi-component)	Social networks All GPs were sent an extract of the latest updated version of the Dutch College of GPs clinical practice guidelines Skills building Mailed guideline + invitation for a 2-hr educational session on asthma and inhalation techniques Multi-component Mailed guideline + educational session + GPs received written individualized treatment advice based on symptoms, the use of medication, lung function, and the severity of AHR.	Treatment of childhood asthma Dutch College of General Practitioner's clinical practice guideline No No	G1: Print G2: Print + in-person G3: Print + in-person + print G1: postal G2: Postal + unclear G3: Postal + unclear + postal All groups: Guideline: 1 session. G2 & G3: Educational session: 1 session; total time: 2 hr	Combination of graphical information (e.g., flow chart) quantitative information.	Outlined decision tree for treatment steps

Table E-4. Key question 1 intervention descriptions (continued)

Author, Year	Groups	Comparators	Clinical Focus of the Evidence, Source of Information, Theory-Based, Prior Testing	Intervention Format, Delivery Agent, Intensity	Evidence Presentation	Message of Intervention
Jain et al., 2006[16]	G1: Passive intervention-guidelines by mail (increase reach) G2: Active intervention (multi-component)	Postal mail/email Mailed a copy of the CPGs to study dietitians Multi-component Local opinion leader + access to website + supporting documents + educational tools +training kits to assist the dietician + posters and pocket cards + audit & feedback + interactive workshop +academic detailing + site reports +reminders + small group session.	Nutrition support; treatment Guideline, Canadian Canadian critical care clinical practice guidelines committee Yes No	G1: paper G2: every type of format G1: postal G2: postal + web + in-person + email, etc NR	NR	G1: Guidelines G2: The different strategies sought to predispose people through awareness, enable them through agreement and adoption, and then reinforce behaviors through adherence.

Table E-4. Key question 1 intervention descriptions (continued)

Author, Year	Groups	Comparators	Clinical Focus of the Evidence, Source of Information, Theory-Based, Prior Testing	Intervention Format, Delivery Agent, Intensity	Evidence Presentation	Message of Intervention
Jousimaa et al., 2002[17]	G1: Computerized version of guidelines (increase ability) G2: Textbook-based version of guidelines (increase reach)	Electronic/digital media Physicians were given the latest CD-ROM version of the guidelines. Those with access to a computer in the consultation room were given a copy of the CD-ROM to be installed on their consultation room computer. If the physicians did not have access to a computer in the consultation room, they were provided with a laptop computer with pre-installed guidelines during the study period. Postal mail/email Physicians were given the latest version of the textbook guidelines. Prior to the study, participating physicians agreed not to use the other version of the guidelines if it was available in the health center, but they could use any other source of information, such as medical journals, books, and colleague consultations.	Primary care; treatment and prevention Guidelines; Evidence-Based Medicine Guidelines (EBMG) No No	G1: CD-ROM or computer-based G2: Paper-based G1: CD-ROM or computer-based G2: Paper-based #: 1 length: used version of guidelines for no less than 4 weeks before data collection total time: NR	NR	Guideline recommendations

Table E-4. Key question 1 intervention descriptions (continued)

Author, Year	Groups	Comparators	Clinical Focus of the Evidence, Source of Information, Theory-Based, Prior Testing	Intervention Format, Delivery Agent, Intensity	Evidence Presentation	Message of Intervention
Junghans et al., 2007[18]	G1: Conventional guideline (increase reach) G2: Ratings about specific patients in vignettes (increase motivation)	Electronic/digital media Participants were provided online guideline paragraphs most relevant to each vignette as well as links to full-text guidelines and detailed information on how the ratings were derived. Opinion leaders Physicians received an electronic prompt to physicians that said "the expert panels recommend __-"	Angina; treatment G1: American heart association and European Society of Cardiology G2: 2 expert panels composed of 5 cardiologists, 1 cardiothoracic surgeon, and 5 general practitioners with an intereste in cardiology No No	Electronic-based Computer 12 vignettes randomly ordered	NR	Physicians read unique vignettes of patients with suspected or confirmed angina based on unique combinations of clinical factors (indications).
Kennedy et al., 2003[19]	G1: Control (not abstracted) G2: Information (increase reach) G3: Interview (increase motivation)	Multi-component video and booklet were sent to women at their home 6 weeks before their consultation Interpersonal outreach G2 + in-person structured interview with a research nurse immediately prior to the initial consultation with their gynecologist.	Menorrhagia; treatment Systematic review of treatment efficacy published in the Effective Health Care series + epidemiological and quality of life surveys on the condition No No	G2: Paper + video G3: Paper + video + in person G2: Postal G3: postal + In person (nurse) G2: 1 session G3: interview length was approx. 30 min. 1 session.	Combined	Booklet emphasized the importance of patient preferences, information about menorrhagia and its causes, treatment options, risks and benefits of surgery, and a personal treatment plan. Video included patient narratives, graphical illustrations, and used color coding to facilitate linkage of the visual material with the information in the booklet.

Table E-4. Key question 1 intervention descriptions (continued)

Author, Year	Groups	Comparators	Clinical Focus of the Evidence, Source of Information, Theory-Based, Prior Testing	Intervention Format, Delivery Agent, Intensity	Evidence Presentation	Message of Intervention
King et al., 2007[20]	G1: Attention control (not abstracted) G2: Counselor via phone (increase motivation) G3: Automated counselor via phone (increase reach)	Usual Care Interpersonal outreach Telephone assisted physical activity counseling by a trained health educator. Telephone contacts were supplemented by informational mailings and pedometer. Electronic/digital media Telephone-assisted physical activity counseling by an automated computer. Telephone contacts were supplemented by informational mailings and pedometer.	Physical activity: Prevention US Department of Health and Human Services Yes Yes	Both: via phone and informational mailings G2: trained health educator G3: automated telephone linked computer system that could 'speak' to participants over the telephone using computer controlled speech generation. Total # of calls completed (M, SD) G2: 13.1 (2.5) G3: 11.8 (4.8) Average call length in minutes G2: 10.7 (5.0) G3: 6.6 (2.2) 10-15 minute sessions	NR: probably qualitative	G2: Health educator offered tailored support and problem solving around physical activity barriers. G2 & G3: The content included physical activity assessment, progress evaluation, individualized problem-solving, goal-setting, feedback, and delivery of positive support and tailored advice. Discussion of cognitive and behavioral strategies, derived from Social Cognitive Theory and the Trans-theoretical Model occurred as appropriate to each person's stage of motivational readiness for change. Provided with informational mailings, pedometer, and PA self-monitoring log.

Table E-4. Key question 1 intervention descriptions (continued)

Author, Year	Groups	Comparators	Clinical Focus of the Evidence, Source of Information, Theory-Based, Prior Testing	Intervention Format, Delivery Agent, Intensity	Evidence Presentation	Message of Intervention
Laprise et al., 2009[21]	G1: CME (increase ability) G2: CME + practice enablers and reinforcers (multi-component)	Skills building CME that was a small-group interactive workshop Champions CME + PER group. nurses visited GPs' offices to implement the clinical intervention. A set of clinical tools was developed to support intervention implementation in the practice.	Cardiovascular: prevention Clinical practice guideline No No	G1: in-person G2: In-person G1: expert cardiologist and GP G2: Trained nurse 2 hrs	NR	CME: included a presentation of the latest CPGs by an expert cardiologist, discussion of 4 cases facilitated by a GP and an interactive response system, and discussion about barriers to guidelines implementation in their practice G2: The following tools were developed: list of target diagnoses; lists of generic and commercial names of all antidiabetic and CV drugs available on the market; decision-making algorithm for chart prompting. The goal of the trained PERs was to address important physician-level practice barriers to prevention: time to screen for at-risk patients, time to search for clinical information in support of decision making, and timely access to experts' recommendations

Table E-4. Key question 1 intervention descriptions (continued)

Author, Year	Groups	Comparators	Clinical Focus of the Evidence, Source of Information, Theory-Based, Prior Testing	Intervention Format, Delivery Agent, Intensity	Evidence Presentation	Message of Intervention
Lien et al., 2007,[22] Svetkey et al., 2003,[23] Young et al., 2009[24]	G1: Advice only (increase reach) G2: Advice + behavioral counseling using established intervention (multi-component) G3: Established intervention + DASH dietary recommendations (multi-component)	Interpersonal outreach Individual session with advice. The interventionist reviews the recommended guidelines, gives advice, and provides printed educational materials, but does not provide behavioral counseling. Multi-component Individual and group sessions with behavioral counseling. Participants are instructed on ways to identify the sodium content of food, to select and substitute lower sodium choices, and to alter sodium content of recipes. Multi-component Individual and group sessions with behavioral counseling + DASH. In addition to the established intervention information in G2, participants in G3 received instruction and counseling regarding DASH dietary recommendations. Participants were asked to monitor intake of fruits,	Hypertension, diet, physical activity, and weight loss; prevention and treatment G1: Guideline; National High Blood Pressure Education Program recommendation G2: JNC-V on Detection. Evaluation, and Treatment of High Blood Pressure G3: JNC-VI recommendations G1: # session: 1 total time: 30 min G2 & G3: # session: 18 face-to-face intervention contacts (14 group meetings and 4 individual counseling sessions). Unclear Unclear	G1: In-person + paper-based G2 & G3: in-person G1: Interventionist (typically a registered dietician) G2 & G3: Nutritionists or health educators	Combined quantitative and qualitative with focus both on social support and understand-ing quantitative guidelines	Review recommended guidelines, give advice, provide counseling G1: Just has advice, no behavioral counseling G2 and G3: include behavioral counseling which sought to increase self-efficacy and outcome expectancies. Message emphasizes the importance of an individual's ability to regulate behavior by setting goals, monitoring progress towards the goals, and attaining skills necessary to reach the goals.

Table E-4. Key question 1 intervention descriptions (continued)

Author, Year	Groups	Comparators	Clinical Focus of the Evidence, Source of Information, Theory-Based, Prior Testing	Intervention Format, Delivery Agent, Intensity	Evidence Presentation	Message of Intervention
Lien et al., 2007,[22] Svetkey et al., 2003,[23] Young et al., 2009[24] (continued)		vegetables, and dairy products in addition to recorded physical activity and calorie and sodium intake.				
Marcus et al., 2009[25]	G1: Contact control treatment delayed group (not abstracted) G2: Telephone-based individualized feedback (increase motivation) G3: Print-based individualized feedback (increase reach)	Multi-component Mailed health education information in the form of tip sheets (stress management, cancer prevention, healthy nutrition and back health) on same schedule as G2 and G3 Telephone Telephone contact with health educator who incorporated feedback generated by the computer expert system and supplemented with stage-based manual and tip sheets but no script. Multi-component Printed reports of feedback generated by the computer expert system along with manuals matched to their stage of motivational readiness for physical activity adoption and tip sheets.	Physical activity CDC/ACSM recommendations Yes Yes	G1: paper-based/mailed G2:telephone, G3: print (not sure if mailed or in-person) G1: postal, G2: telephone [health educator], G3: NR #:14 contacts for each participant in each intervention group (G2 and G3) length: For G2: 13 minutes mean call time G3: NR total time: 182 minutes for G2: NR for G3	NR - materials available on request from authors	Individually tailored messages generated by a computer expert system, stage-targeted booklets and physical activity-related tip sheets to both groups. Participants in both treatment conditions (G2 and G3) were instructed that their goal was to increase their moderate intensity physical activity to a level that met or exceeded CDC/ACSM recommendations (at least 5 days per week for a total of at least 30 minutes per day)

Table E-4. Key question 1 intervention descriptions (continued)

Author, Year	Groups	Comparators	Clinical Focus of the Evidence, Source of Information, Theory-Based, Prior Testing	Intervention Format, Delivery Agent, Intensity	Evidence Presentation	Message of Intervention
Maxwell et al.,2010[26]	G1: Control (not abstracted) G2: Educational session + letter to provider (multi-component) G3: Educational session + letter to provider + FOBT kit (multi-component)	Multi-component educational session facilitated by a trained Filipino American health educator Multi-component Educational session facilitated by a trained Filipino American health educator: PLUS free FOBT kit and asked to sign pledge that they would use it	Colorectal cancer screening American Cancer Society and Task Force on Community Preventive Services Yes Yes	Paper-based and in-person Delivered in community-based networks by a trained nurse #:1 length: 60-90 minutes total time	Combined	Discussed CRC and screening information, demonstrated use of FOBT kit, and addressed barriers to screening as well as peer-to-peer feedback, also received print materials and final RN recommendation to get screened

Table E-4. Key question 1 intervention descriptions (continued)

Author, Year	Groups	Comparators	Clinical Focus of the Evidence, Source of Information, Theory-Based, Prior Testing	Intervention Format, Delivery Agent, Intensity	Evidence Presentation	Message of Intervention
Murtaugh et al.,2005[27]	G1: Usual care (not abstracted) G2: Basic intervention e-mail reminder (increase reach) G3: Augmented intervention of e-mail reminder + package of supporting materials (multi-component)	Postal mail/email "just-in-time" e-mail reminder on an initial screen listing the 6 key CHF practices in very abbreviated form (spelling acronym "ADHERE"). Subsequent links with more detailed information. Multi-component "Just-in-time" e-mail reminder like comparator 1 plus package of material stating it was for the care of the CHF pt (including laminated pocket card on med management, prompter card to improve communication with MDs, and self-care guide for patients); also received "expert peer" follow-up outreach including follow-up email reminder, inquiries on the usefulness of the card distributed within 10 days of admission to home care.	Heart failure management NR Yes Unclear	Electronic-based and paper-based E-mail and paper-based, but e-mail sent by "expert peer" #: NR length: within 45 days of pt admission total time: NR	NR	Quick reference sheet and more detailed recommendations on medications, documenting vital signs and symptoms/signs; pt records daily weights, education about low sodium, heart failure symptoms, and education about the "heart failure self-care" guide.

Table E-4. Key question 1 intervention descriptions (continued)

Author, Year	Groups	Comparators	Clinical Focus of the Evidence, Source of Information, Theory-Based, Prior Testing	Intervention Format, Delivery Agent, Intensity	Evidence Presentation	Message of Intervention
Paradis et al.,2011[28]	G1: Paper handouts (increase reach) G2: Educational DVD (increase reach)	Usual Care A packet of written handout materials already available in clinic that covered similar (though not identical) information to that shown in the video (comparator #2). All written handouts were at a fourth-grade readability level. Families could take these materials home with them. Electronic/digital media A locally produced DVD that depicted basic aspects of newborn care. Topics covered included normal newborn breathing patterns, bathing and feeding, safe sleeping practices, dealing with crying, and promoting development. A local pediatrician and several ethnically diverse babies appeared in the video. After viewing the video in the clinic, families were given the video to take home with them.	Newborn care; prevention and management Guidelines; American Academy of Pediatrics No No	G1: Paper-based G2: Video G1 and G2 were delivered in the clinic by a staff member #: 1 length: 15 minutes total time: 15 minutes	NR	Basic aspects of Newborn care as depicted by guidelines. Topics included normal newborn breathing patterns, bathing and feeding, safe sleeping practices, dealing with crying, and promoting development.

Table E-4. Key question 1 intervention descriptions (continued)

Author, Year	Groups	Comparators	Clinical Focus of the Evidence, Source of Information, Theory-Based, Prior Testing	Intervention Format, Delivery Agent, Intensity	Evidence Presentation	Message of Intervention
Partin et al., 2004[29]	G1: Usual care (not abstracted) G2: Pamphlet (increase reach) G3: Video (increase reach)	Usual Care Note: Usual care never described so it was not abstracted. Postal mail/email Pamphlet that provided a balanced representation of the potential risks and benefits of screening. Electronic/digital media Video designed to provide a balanced representation of the risks and benefits of screening	Prostate cancer screening G2: unclear G3: Foundation for Informed Medical Decision Making Yes No	G2: Paper-based G3: Video G2: Postal G3: Postal. (In the video two physicians (an internist and urologist) and patient delivered information) G2: 1 time exposure, 8 page pamphlet G3: 23 min video, 1 time exposure	G2: Qualitative G3: Qualitative and graphical	G2: Written at 6th grade level. It starts with a definition of the PSA and why not all doctors are recommending it. It defines the prostate and CaP and how CaP is different from the common but less serious condition, BPH, which causes similar symptoms. It then summarizes the accuracy of the PSA and the unknown efficacy of CaP treatments. Space is provided on the back to write down questions to discuss with a health care provider. The point that there is a decision to make and that the patient should play an active role in it is emphasized throughout. G3: Designed to enable 100% comprehension at the 10th grade level. Video developed by the Foundation for Informed Medical Decision Making. It uses physician actors who articulate the advantages and disadvantages of testing,

Table E-4. Key question 1 intervention descriptions (continued)

Author, Year	Groups	Comparators	Clinical Focus of the Evidence, Source of Information, Theory-Based, Prior Testing	Intervention Format, Delivery Agent, Intensity	Evidence Presentation	Message of Intervention
Partin et al., 2004[29] (continued)						presents testimonials from patients, and shows graphic illustrations to promote an informed decision. Viewers are asked to consider 3 questions in making a decision about screening and are encouraged to discuss screening with their doctors.

Table E-4. Key question 1 intervention descriptions (continued)

Author, Year	Groups	Comparators	Clinical Focus of the Evidence, Source of Information, Theory-Based, Prior Testing	Intervention Format, Delivery Agent, Intensity	Evidence Presentation	Message of Intervention
Rahme et al., 2005[30]	G1: No treatment control (not abstracted) G2: Decision tree (increase ability) G3: Workshop (increase ability) G4: Workshop + decision tree (multi-component)	Additional resources A laminated sheet representing the decision tree was distributed to physicians in the decision tree group, followed by a letter of explanation from the Continuing Medical Education Department regarding the content and use of the decision tree, without any further justification or discussion of the medical content. Skills building Small-group 90 minute workshops modeled after the Script Concordance test. The decision tree was presented during the workshops for the workshop group but the laminated sheet was not distributed Multi-component Workshop + laminated decision tree	Osteoarthritis treatment Evidence-based clinical practice guidelines - American College of Rheumotology No No	G2: Paper-based G3: In-person G4: Paper-based +in-person G2: In-person by sales representatives G3/G4: In - person (peer-facilitated by a general practitioner and a rheumatologist who served as a resource person) G3/G4: 90 minute workshop	G2: Qualitative G3/G4: Unclear	The decision tree discussed treatment choices for osteoarthritis patients, suggesting nonpharmacological treatment including physical exercise as first-line therapy, and pharmacological choices starting with acetaminophen and moving to NSAIDs or COX-2 inhibitors with or without a gastroprotective agent, depending on the patient response to treatment and the presence of risk factors for NSAID gastropathy. The workshop discussed evidence-based management of patients with osteoarthritis.

Table E-4. Key question 1 intervention descriptions (continued)

Author, Year	Groups	Comparators	Clinical Focus of the Evidence, Source of Information, Theory-Based, Prior Testing	Intervention Format, Delivery Agent, Intensity	Evidence Presentation	Message of Intervention
Rebbeck et al., 2006[31]	G1: Dissemination of guidelines by mail (increase reach) G2: Implementation group (multi-component)	Postal mail/email Dissemination of guidelines by mail Skills building A one-day (8 hour) workshop, which included interactive sessions outlining the content of the guidelines, practical sessions covering the treatments endorsed in the guidelines (i.e., 'reassure patient' and 'advise to act as usual'), and the use of functional outcome measures. Physiotherapists also given a laminated copy of the algorithms outlining the process of care, appointment cards, and marketing material to be used for general practitioners who usually refer to the practice. A follow-up educational outreach visit (2 hrs) approximately 6 months later, involving problem solving regarding use of the guidelines in clinical practice and an update of the evidence given.	Whiplash treatment/ management Clinical practice guidelines, developed by the MAA No No	G1: paper-based G2: in-person G1: postal mail G2: in-person, delivered in part by opinion leader G2: Educational intervention #: 2 length: 8 hours and 2 hours total time: 10 hrs.	NR	G1: guideline recommendations G2: guideline recommendations + information and help with problem solving

Table E-4. Key question 1 intervention descriptions (continued)

Author, Year	Groups	Comparators	Clinical Focus of the Evidence, Source of Information, Theory-Based, Prior Testing	Intervention Format, Delivery Agent, Intensity	Evidence Presentation	Message of Intervention
Rimer et al., 2001[32]	G1: No treatment control/usual care (not abstracted) G2: TP (increase reach) G3: TC (multi-component)	Postal mail/email TP: "(PRISM; 7x9 in, full-color booklet with graphic images included tailored colored pie charts to illustrate risk-related information; Addressed personally to recipient with tailoring "especially for you", based on prior interview's info-pt's stage of readiness; also had personal risk of breast cancer in next 10 years using Gail model; overall women's risks by age group, etc. Tailored on 20 items Multi-component TP +TC- same as TP + a call by trained advisors asking open-ended questions about the booklet to elicit discussion about breast cancer and mammography; discussed Gail scores, addressed barrier to screening and other concerns; communicated guidelines	Breast cancer screening NIH Consensus Conference on Breast Cancer Screening Yes Yes	Paper-based and telephone Postal or telephone (telephone was by research staff) TP #: 1 length: 20-25 pages total time: NR TP + TC #: 1 length: none total time: NR	Combined	Guidelines, tailored statistics, risk factors, barriers

Table E-4. Key question 1 intervention descriptions (continued)

Author, Year	Groups	Comparators	Clinical Focus of the Evidence, Source of Information, Theory-Based, Prior Testing	Intervention Format, Delivery Agent, Intensity	Evidence Presentation	Message of Intervention
Simon et al., 2005[33]	G1: Mailed educational materials (increase reach) G2: Individual academic detailing (increase ability) G3: Group academic detailing (increase ability)	Postal mail/email Mailing that contained printed material describing the current guidelines for prescribing antihypertensive medications and a laminated wallet card that summarized the guidelines Skills building Mailing (same as comparator 1) + one-on-one educational outreach meetings which consisted of a single visit (15-30 minutes) from the trained detailer, incorporating the core principles and methods of academic detailing Skills building Mailing (same as comparator 1) + 45-minute small-group academic detailing sessions; also employed supportive group processes, such as encouraging clinicians to share success stories in overcoming barriers to adhering to guideline recommendations and providing clinicians with an opportunity for mutual reinforcement of desired practice behaviors.	Hypertension; treatment Guidelines No No	G1: Paper-based G2: In-person G3: In-person G1: Postal mail G2: In-person, delivered by respected physician idea champion G3: in-person via group, delivered by respected physician idea champion G2: #: 1 length: 15-30 minutes total time: 15-30 min. G3: #: 1 length: 45 minutes total time: 45 min.	G1: Unclear G2: Combined G3: Combined	Academic detailing; guideline recommendations

Table E-4. Key question 1 intervention descriptions (continued)

Author, Year	Groups	Comparators	Clinical Focus of the Evidence, Source of Information, Theory-Based, Prior Testing	Intervention Format, Delivery Agent, Intensity	Evidence Presentation	Message of Intervention
Soler et al., 2010[34]	G1: Control (not abstracted) G2: Training session on the SEPAR guidelines (increase ability) G3: G2 + portable-device for spirometry (multi-component)	Skills building G2: GPs dealing with COPD received training session based on the literal transcription of the SEPAR-SEMYC guidelines for the diagnosis, severity stratification and management of COPD. Training was performed by pulmonologists from Spanish hospital institutions who had previous information about the SEPAR guidelines Multi-component G3: G2 intervention plus the GPS in G3 attended a spirometry training session on the KoKo Peak Pro devices immediately after the SEPAR guidelines presentation.	COPD Chart Unclear No	Paper-based, in-person In person #: 1 training session for participants in G2 and G3 length: NR total time: NR	Unclear	NR

Table E-4. Key question 1 intervention descriptions (continued)

Author, Year	Groups	Comparators	Clinical Focus of the Evidence, Source of Information, Theory-Based, Prior Testing	Intervention Format, Delivery Agent, Intensity	Evidence Presentation	Message of Intervention
Sullivan et al., 2010[35]	G1: VA guidelines (increase reach) G2: COPE: web-based education program (increase ability)	Postal mail/email Residents accessed online through links in email and completed training individually, deciding how much time to spend on it. VA guidelines are a text document that uses a modular approach "to provide a scientific evidence base for practice interventions and evaluations, specifically in the use of opioids to treat CNCP." It contains clinical algorithms clinicians can use to "determine the best interventions and timing of care for their patients, reduce the incidence of adverse-effects and other undesirable outcomes, and optimize healthcare utilization." Key points and a treatment algorithm flow chart provide a distillation of the recommendations Skills building Residents accessed online through linksin email and completed training individually. The COPE training focuses on communication challenges	Chronic non-cancer pain; treatment Guidelines; Veterans Affairs/Dept. of Defense No No	G1 and G2 were both web-based E-mail with links to intervention for G1 and G2 G1: #: 1 length: 26 chapters total time: up to individual G2: #: 1 length: 6 chapters total time: up to individual	Unclear NOTE: Over 100 web pages depict clinical interactions between simulated physicians and patients with supporting scientific, policy, and clinical material. Basic factual material about opioid pharmacology, opioid effectiveness for CNCP, and the risks of chronic opioid therapy are presented in the first chapter. Depicts interactions with one patient at low-risk for poor outcome from opioid therapy and one patient at high risk for poor outcome. A summary chapter provides take home points and	G1: Guideline recommendations G2: Skill-building and help with problem solving in shared decision making for cancer treatment

Table E-4. Key question 1 intervention descriptions (continued)

Author, Year	Groups	Comparators	Clinical Focus of the Evidence, Source of Information, Theory-Based, Prior Testing	Intervention Format, Delivery Agent, Intensity	Evidence Presentation	Message of Intervention
Sullivan et al., 2010[35] (continued)		between physicians and patients with CNCP who are using long-term prescription opioids. Presents a shared decision-making procedure. Over 100 web pages depict clinical interactions between simulated physicians and patients with supporting scientific, policy, and clinical material. Basic factual material about opioid pharmacology, opioid effectiveness for CNCP, and the risks of chronic opioid therapy are presented in the first chapter. Depicts interactions with one patient at low-risk for poor outcome from opioid therapy and one patient at high risk for poor outcome. A summary chapter provides take home points and prinTable E-Patient Treatment Agreements, Survival Tips, and key Helpful Phrases to use with patients. Interactive quizzes engage the viewer in clinical problem solving.			prinTable E-Patient Treatment Agreements, Survival Tips, and key Helpful Phrases to use with patients. Interactive quizzes engage the viewer in clinical problem solving.	

Table E-4. Key question 1 intervention descriptions (continued)

Author, Year	Groups	Comparators	Clinical Focus of the Evidence, Source of Information, Theory-Based, Prior Testing	Intervention Format, Delivery Agent, Intensity	Evidence Presentation	Message of Intervention
Watson et al., 2002[36]	G1: Guideline materials by postal mail (increase reach) G2: EO session and guidelines (increase ability) G3: CPE session and guidelines (increase ability) G4: Guidelines + EO and CPE (multi-component)	Postal mail/email Guideline materials mailed to all pharmacies in the Grampian region of Scotland Interpersonal outreach One outreach visit by a trained pharmacist and a followup phone call 4-6 weeks later to determine whether the guidelines were being used and whether there had been any problems or queries with their use Skills building Invitations to attend one of three CPE sessions arranged at different venues; each session followed a standard SCPPE format and comprised a 1 hour presentation on vulvovaginal candidiasis by ta consultant or genito-urinary medicine; a 90 minute case study workshop and practice applying guidelines. CPE occurred prior to outreach visit	OTC management of vulvovaginal candidiasis Cochrane Review (2001) by the same authors of this study. Title: Oral versus intra-vaginal imidazole and trazole anti-fungal treatment of acute, uncomplicated vulvovaginal candidiasis Yes Unclear	Paper-based, in-person Postal, pharmacy-based, in-person by pharmacist G2: One visit and 1 followup phone call at 4-6 weeks G3: CPE session: 1 hour presentation, 90 minute case study workshop, total time: 2.5 hours G4: G2+G3	Combined	Guideline recommendations

Table E-4. Key question 1 intervention descriptions (continued)

Author, Year	Groups	Comparators	Clinical Focus of the Evidence, Source of Information, Theory-Based, Prior Testing	Intervention Format, Delivery Agent, Intensity	Evidence Presentation	Message of Intervention
Wetter et al., 2006[37]	G1: Single standard telephone-counseling session (increase reach) G2: Multiple enhanced telephone counseling sessions (multi-component)	Skills building SC Consisted of the single CIS counseling session that had been delivered during the initial call to the CIS, plus an offer of Spanish language self-help materials that would be mailed to the participant. Skills building Enhanced Counseling was Standard counseling plus 3 additional proactive counseling calls; involved practical counseling (the identification of triggers to smoke and high risk situations, as well as coping strategies); social support from counselor and assisting participant in strategies for obtaining social support in their environment; motivational enhancement techniques; culturally tailored	Smoking cessation; prevention Guideline Yes No	Phone-based supplemented by printed materials Counselors from CIS and research team G1: #: 1 call length: NR total time: NR G2: #: 4 calls length: Call 2: M=16 min; Call 3: M=15 min; Call 4: M=14 min total time: Approx. 45 min + initial call	Qualitative	Motivational, coping, social support

Table E-4. Key question 1 intervention descriptions (continued)

Author, Year	Groups	Comparators	Clinical Focus of the Evidence, Source of Information, Theory-Based, Prior Testing	Intervention Format, Delivery Agent, Intensity	Evidence Presentation	Message of Intervention
Wolters et al., 2005[38]	G1: Control mailed guidelines (increase reach) G2: Intervention involving package for learning, supporting materials, decision tree, and information leaflets for patients (multi-component)	Postal mail/email Sent existing national guidelines on LUTS Multi-component Package for learning + supporting materials + decision trees + information leaflets for patients	LUTS Dutch College of General Practitioner's clinical practice guideline No Unclear	Paper-based Postal 1 time	NR	G2: Items designed to enhance knowledge. PIL contained background information, package of questions reflecting on a recent male patient attending surgery with LUTS, the clinical management of hypothetical four cases, clinical management of LUTS, statements about (fear of) prostate cancer, and possible barriers around bladder catheterization in care of acute urinary retention. The consultation supporting materials included Dutch College of General Practitioners guidelines on Lower urinary Tract Symptoms summarized on a A5 format card, The guideline summarized in two decision trees, IPSS, BS, Voiding dairy. The patient information leaflet talked about the causes of LUTS and treatment options and prostate carcinoma in relation to LUTS and the limitations of PSA-testing

Table E-4. Key question 1 intervention descriptions (continued)

Author, Year	Groups	Comparators	Clinical Focus of the Evidence, Source of Information, Theory-Based, Prior Testing	Intervention Format, Delivery Agent, Intensity	Evidence Presentation	Message of Intervention
Wright et al., 2008[39]	G1: Standardized lecture by expert opinion leader (increase motivation) G2: Standardized lecture by expert opinion leader + academic detailing and a toolkit (multi-component)	Opinion leaders Standardized formal lecture led by expert opinion leader in colon cancer. The lecture emphasized the importance of adequate lymph node assessment in colon cancer management to the local surgeons and pathologists. Multi-component Standardized formal lecture led by expert opinion leader in colon cancer (as in comparator 1); also, the expert opinion leader met with locally identified opinion leaders in colon cancer to discuss the importance of adequate lymph node assessment, local barriers to improving lymph node assessment, and possible solutions (academic detailing) and provided the local opinion leader with a toolkit containing a pathology template and a poster and pocket cards that emphasized that 12 lymph nodes should be assessed. A follow-up reminder package was sent 6 months after the presentation to the treatment	Colon cancer; treatment Guidelines No No	in-person Expert and local opinion leaders in colon cancer G1: Expert opinion leader in colon cancer G2: Expert and local opinion leaders in colon cancer G1: #: 1 lecture length: NR total time: NR G2: #: 1 lecture + one academic detailing session length: 15-30 minutes total time: NR	NR	Guideline recommendations

Table E-4. Key question 1 intervention descriptions (continued)

Author, Year	Groups	Comparators	Clinical Focus of the Evidence, Source of Information, Theory-Based, Prior Testing	Intervention Format, Delivery Agent, Intensity	Evidence Presentation	Message of Intervention
Wright et al., 2008[39] (continued)		group only, which included a cover letter from the expert opinion leader, a peer-reviewed article regarding optimization of lymph node assessment by using lymph node clearing solutions, and more of the same pocket cards.				

Abbreviations: ADHERE = acronym for six key heart failure clinical practices for improved patient health outcomes; AF = audit and feedback; AHCPR = Agency for Health Care Policy and Research; AHR = airway hyper-responsiveness; ATP III = Adult Treatment Panel III; BPH = benign prostatic hypertrophy; BS = Bother score; CAL = computer-assisted learning; CaP = Cancer of the Prostate; CDC/ACSM = Centers for Disease Control and American College of Sports Medicine (joint study); CD-ROM = pre-pressed compact disc that contains data accessible to, but not wriTable E-by, a computer for data storage and music playback; CHF = congestive heart failure; CIS = Computer Information Service; CME = continuing medical education; CNCP = Chronic non-cancer pain; COPD = chronic obstructive pulmonary disease; COPE = Compassionate Options for Progressive Eldercare; COX-2 = Cyclooxygenase-2; CPG = Clinical Practice Guideline; CRC = colorectal cancer; CV = cardiovascular; DASH = Dietary Approaches to Stop Hypertension; DEGAM = German College of General Practitioners and Family Physicians; DVD = optical disc storage format; EMR = electronic medical record; FOBT = fecal occult blood test; G = group; GP = general practitioner; hr = hour; IPSS = International Prostate Symptom Score; JNC-V = Joint National Committee; KNGF = Royal Dutch Society for Physical Therapy; LBP = lower back pain; LHA = lay health advisor; LUTS = lower urinary tract symptoms; MAA = Motor Accidents Authority; NIH = National Institutes of Health; NR = not reported; NSAID = nonsteroidal anti-inflammatory drug; OTC = Over the counter; PA = physician's assistant; PDA = personal digital assistant; PER = practice enablers and reinforcers; PIL = packaged for individual learning; PRISM = Personally Relevant Information about Screening mammography; PSA = prostate-specific antigen; pt = patient; Q&A = Questions and Answers; SC = Standard Counseling; SCPPE = _; SIGN = Scottish intercollegiate guideline network; TP = Tailored print; TC = Tailored print and telephone counseling; TPV = tailored and targeted print and video; US = United States; USPTF = US Preventive Services Task Force; VA = Veterans Administration.

Table E-5. Key question 2 studies first outcome

Author, Year	Groups	Outcome #1, Exact Measure Used	Timing of Measurement, Data Source	N analyzed for This Outcome	Results by Group	Differences in Groups	Covariates Controlled for in Analysis, Statistical Methods Used
Bahrami et al., 2004[1]	G1: Mailed guideline (increase reach) G2: Guideline + AF (not abstracted) G3: CAL (increase ability) G4: CAL + AF (not abstracted)	Behavior (applicable for clinicians) Proportion of patients whose treatment complied with the guideline. Assessed by two independent researchers and any disagreements were resolved by discussion.	4 month period in 1999 (pre-intervention) 4 month period in 2000 (post-intervention) Clinical records	Patients Pre-: 3342 Post: 1934 Dentists G1: 11 G3: 11	Pre-intervention % (95%CI) G1: 77% (70/85%) G3: 70% (56/84%) Post-intervention G1:81% (70-92%) G3: 73% (59-88%)	NR :	Pericoronitis, caries and pulpal pathology Weighted t-test
Banait et al., 2003[2]	G1: Mailed guidelines (increase reach) G2: Educational outreach (Multi-component)	Behavior (applicable for clinicians) Appropriateness of referrals for open access endoscopy. Proportion of appropriate referrals. Referrals for open access endoscopy were included if the GP had requested the procedure without a prior hospital consultation. The characteristics of each referral made in the 7 months following the initial outreach visit were appraised using predefined medical review criteria based on the guidelines.	7 months following outreach visit Chart	G1: 36 G2 (ITT): 44 G2 (only those that accepted invitation to participate in intervention): 27	Median percentage of appropriate referrals per practice (IQR) G1: 50.0 (221./72.4) G2 (ITT): 63.9 (50.0/100.00) G2: 72.7 (50.0/100.0)	Difference between control and intervention practices: Mann-Whitney z: -2.235, 1 df, p=0.025	Used when appropriate", but doesn't provide more details. Non-parametric tests

Table E-5. Key question 2 studies first outcome (continued)

Author, Year	Groups	Outcome #1, Exact Measure Used	Timing of Measurement, Data Source	N analyzed for This Outcome	Results by Group	Differences in Groups	Covariates Controlled for in Analysis, Statistical Methods Used
Beaulieu et al., 2004[3]	G1: Control (not abstracted) G2: Guideline (increase reach) G3: Guideline + reminder notice and stickers for patients' charts (multi-component)	Behavior (applicable for clinicians) Treatment of sTable E-angina in line with guideline. Measured by looking at the prescription of 3 cardiovascular medications in 1999. Data are odds ratios (95%CI) for receiving a prescription for the class of drug	6 months post intervention Computerized database of the Quebec health insurance board	Total: 2326 G2: 766 G3: 793	β-Blocker G2: 1.00 (0.88/1.13) G3: 1.04 (0.92/1.18) Antiplatelet G2: 1.05 (0.94/1.18) G3: 1.07 (0.95/1.20) Hypolipaemics G2: 1.02 (0.90/1.16) G3: 0.95 (0.83/1.08)	β-Blocker G2 vs. G3: 0.04 p=NR Antiplatelet G2 vs. G3: 0.02 p=NR Hypolipaemics G2 vs. G3: 0.07 p=NR	Took into account covariance between observations sharing the same hierarchical structure multilevel logistic regression
Becker et al., 2008[4]	G1: Mailed guideline (Increase clinician reach) G2: Guideline implementation (multi-component, clinicians only) G3: Guideline implementation and motivational counseling directed at patient (multi-component, clinicians and patients)	Clinical outcomes (applicable for general public/patients) Functional capacity Hannover Functional Ability Questionnaire for Measuring Back Pain-Related Functional Limitations. Normal function shows scores of 80% to 100%, scores around 70% equal a moderately, scores below 60% a severely limited function. Days in Pain	Baseline and at 6 months and at 12 months Self-administered questionnaire	Patient N baseline = 1378 G1: 479 G2: 489 G3: 410 N 6 months = 1261 G1: 450 G2: 435 G3: 376 N 12 months = 1211 G1: 425 G2: 421 G3: 365	Functional capacity: 6 months G1: M = 70.29 G2: M = 72.94 G3: M = 73.94 12 months G1: M = 71.56 G2: M = 72.96 G3: M = 74.64 Days in pain 6 months G1: M = 80.78 G2: M = 63.35 G3: M = 62.91 12 months G1: M = 71.32 G2: M = 58.48 G3: M = 61.57	Functional capacity (odds ratios for groups compared with control only) 6 months Mean diff (95% CI) G1 vs. G2: 2.65 (-0.70/6.01) G1 vs. G3: 3.65 (0.32/6.98) G2 vs. G3: 0.999* p=NR 12 months Mean diff (95% CI) G1 vs. G2: 1.40 (-2.24/5.02) G1 vs. G3: 3.11 (-0.47/6.70) G2 vs. G3: 1.681* p=NR Days in Pain	Sex, age, fear avoidance, physical activity, and number of days in pain during previous 6 months Multilevel mixed modeling accounting for clustering of data on practice level

Table E-5. Key question 2 studies first outcome (continued)

Author, Year	Groups	Outcome #1, Exact Measure Used	Timing of Measurement, Data Source	N analyzed for This Outcome	Results by Group	Differences in Groups	Covariates Controlled for in Analysis, Statistical Methods Used
Becker et al., 2008[4] (continued)						6 months G1 vs. G2: -16.43 (-26.83/-6.03) G1 vs. G3: -17.87 (-28.18/-7.55) G2 vs. G3: 0.434* p=NR 12 months G1 vs. G2: -12.84 (-23.38/-2.30) G1 vs. G3: -9.76 (-20.20/-0.69) G2 vs. G3: 3.085* p=NR	
Bekkering et al., 2005[5]	G1: Received guidelines by mail (increase reach) G2: Received guidelines + active training strategy (multi-component)	Behavior (applicable for clinicians) Adherence to 4 recommendations. Proportion of patients for whom each and all 4 were fulfilled.	Baseline and followup (exact time not specified) Chart	physiotherapists G1: 48 G2: 37 Patients G1: 253 G2: 247	Limit # of sessions in normal course: G1: 13% G2: 27% Set functional treatment goals G1: 71% G2: 79% Use mainly active interventions G1: 605 G2:77% Give adequate information G1: 87% G2: 96% All four recommendations G1: 30% G2: 42%	Effect of strategy OR (95%CI) Limit # of sessions: 2.39 (1.12/5.12) Set functional treatment goals 1.99 (1.06/3.72) Use mainly active interventions 2.79 (1.19/6.55) Give adequate interventions 3.59 (1.35/9.55) All four 2.05 (1.15/3.65)	Postgraduate education in low back pain logistical multilevel analyses

Table E-5. Key question 2 studies first outcome (continued)

Author, Year	Groups	Outcome #1, Exact Measure Used	Timing of Measurement, Data Source	N analyzed for This Outcome	Results by Group	Differences in Groups	Covariates Controlled for in Analysis, Statistical Methods Used
Bishop and Wing, 2006[6]	G1: Control (not abstracted) G2: Physician only (increase reach) G3: Physician and patient (multi-component)	Behavior (applicable for clinicians) Guideline-concordant treatment advice for 0-4 week post onset. The compulsory WCB physician report forms were collected and scored. Dichotomous measure of 1 = presence of concordant/ Discordant behavior.	Once at 0-4 weeks WCB reports	0-4 weeks Overall: 462 G2: 162 G3: 151	Concordant Behavior: Education & Reassurance G2: 10% G3: 6% Exercise: G2: 38% G3: 53% Appropriate Medication: G2: 85% G3: 81% Spinal Manipulation G2: 2.5% G3: 5% Discordant Behavior: Bedrest: G2: 10% G3: 18% NOTE: Authors did not provide any figures, tables, or data for the >12 week measures. Only state no change seen in the recommended use of ongoing supervised exercise programs.	Percentage difference (authors only compared groups with control) Education & Reassurance: G2 vs. G3: 4%*, p=NR Exercise: G2 vs. G3: 15%* p=NR G1 vs. G3: 10% difference, p=0.05 Appropriate Medication G2 vs. G3: 4%*, p=NR Spinal Manipulation G2 vs. G3: 2.5%*, p=NR Bedrest G2 vs. G3: 8%*, p=NR Control vs. G2: p=0.05 NOTE: Authors did not analyze between groups difference from each other. Only state no change seen in the recommended use of ongoing supervised exercise programs.	None Chi-square

Table E-5. Key question 2 studies first outcome (continued)

Author, Year	Groups	Outcome #1, Exact Measure Used	Timing of Measurement, Data Source	N analyzed for This Outcome	Results by Group	Differences in Groups	Covariates Controlled for in Analysis, Statistical Methods Used
Bishop and Wing, 2006[6] (continued)						Appears all p-values apply to comparisons with the control group, not among G2 and G3. Bedrest data are for 5-12 weeks, while other data are for 0-4 weeks.	

Table E-5. Key question 2 studies first outcome (continued)

Author, Year	Groups	Outcome #1, Exact Measure Used	Timing of Measurement, Data Source	N analyzed for This Outcome	Results by Group	Differences in Groups	Covariates Controlled for in Analysis, Statistical Methods Used
Campbell et al., 2004[7]	G1: Control (not abstracted) G2: LHA (increase motivation) G3: TPV (multi-component) G4: TPV and LHA (multi-component)	Health-related decisions or behavior (applicable for general public/patients) Diet. Dietary fruit and vegeTable E-consumption were measured with the 60-item version of the National cancer health habits and history food frequency questionnaire. The questionnaire assesses frequency of consumption and portion size. The Block database was then used to determine fat consumption, percentage of calories from fat, and number of daily servings of fruits and vegetables. Results shown as servings per day (Mean, Standard Error)	Baseline and 1 yr follow-up Self-report	N = 587 G2: 123 G3: 159 G4: 176	Fruit and vegeTable E-servings/day Baseline G2: 3.5 (0.18) G3: 3.3 (0.16) G4: 3.4 (0.15) Follow-up G2: 3.5 (0.18) G3: 3.9 (0.16) G4: 3.7 (0.15) % meeting 5-a-day recommendations baseline G2: 16.0 G3: 18.9 G4: 19.5 Follow-up G2: 15.4 G3: 21.7 G4: 26.4	G2 vs. G3: 0.2 G2 vs. G4:0.1 G3 vs. G4: 0.1 ns p=0.87 Follow-up G2 vs. G3: 0.4 G2 vs. G4:0.2 G3 vs. G4: 0.2 p=0 .02 for the TPV "intervention main effect" (NOTE: believe meaning the main effect from the TPV/LHA interaction term, but the main effect is compared to control group in all cases in this study) % meeting 5-a-day recommendations baseline G2 vs. G3: 2.9 G2 vs. G4: 3.5 G3 vs. G4: 0.6 ns, p=0 .34 follow-up G2 vs. G3: 6.3 G2 vs. G4: 11.0 G3 vs. G4: 4.7 p=0.04 for the TPV "intervention main effect" (see above)	Demographics Regression models

Table E-5. Key question 2 studies first outcome (continued)

Author, Year	Groups	Outcome #1, Exact Measure Used	Timing of Measurement, Data Source	N analyzed for This Outcome	Results by Group	Differences in Groups	Covariates Controlled for in Analysis, Statistical Methods Used
Carney et al., 2005[8]	G1: Mailed health information (increase reach) G2: Telephone counseling (increase motivation)	Health-related decisions or behavior (applicable for general public/patients) Adherence to screening. To determine participants‘ levels of adherence to screening, the dates of all mammographic encounters that occurred among women in the study were entered into the analysis database. Coded as dichotomy	Over the span of a year Objective measurement; NIH mammography registry	Overall N: 258 G1: 126 G2: 132	Between 1st and 2nd intervention: G1: 47.7% G2: 60.3% Between 15 months and after 2nd intervention: G1: 34.8% G2: 41.3%	Difference in groups between 1st and 2nd intervention: 12.6%, p=0.04 Difference in groups between 15 months and after 2nd intervention: 6.5%, p=0.29	NR Chi-square

Table E-5. Key question 2 studies first outcome (continued)

Author, Year	Groups	Outcome #1, Exact Measure Used	Timing of Measurement, Data Source	N analyzed for This Outcome	Results by Group	Differences in Groups	Covariates Controlled for in Analysis, Statistical Methods Used
Christakis et al., 2006[9]	G1: Usual care (not abstracted) G2: Parental content Alone (increase reach) G3: Provider notification alone (not abstracted) G4: Parental content and provider notification (multi-component)	Discussions about the evidence "At your child's most recent checkup on [date of last visit], did you and your child's doctor discuss [each topic]?" All parents were asked about all of the relevant prevention topics targeted by MyHealthyChild, regardless if they had expressed interest.	2 to 4 weeks after scheduled well-child visit, participants completed a telephone interview Self-report	Unclear	IRR (95%CI) G2: 1.05 (0.97-1.13) G4: 1.09 (1.00-1.20)	G2 vs. G4: 0.04*	NR Poisson analysis

Table E-5. Key question 2 studies first outcome (continued)

Author, Year	Groups	Outcome #1, Exact Measure Used	Timing of Measurement, Data Source	N analyzed for This Outcome	Results by Group	Differences in Groups	Covariates Controlled for in Analysis, Statistical Methods Used
Davis et al., 2004[10]	G1: Control - guidelines by mail (increase reach) G2: Intermediate (multi-component) G3: High intervention (multi-component)	Clinical outcomes (applicable for general public/patients) SF-36 general health-related quality of life instrument. Mean composite scores range from 0-100. Higher scores represent better patient-perceived health related QOL.	baseline and 12 month follow-up Self-report	Patients at Baseline: Overall:1,133 G1: 370 G2: 364 G3: 399 Patients at follow-up Overall: 811 G1: 255 G2: 269 G3: 287	Baseline scores with 95% CI Mental component summary G1: 47.7 (45.2/50.2) G2: 49.7 (48.1/51.3) G3: 49.8 (47.9/51.7) Physical component summary G1: 44.4 (42.5/46.2) G2: 45.8 (43.2/48.4) G3: 43.6 (41.5/45.6) General health profile G1: 63.7 (58.3/69.2) G2: 67.6 (64.9/70.3) G3: 62.1 (59.1/65.1) 12 month follow-up score with 95% CI Mental component summary: G1: 48 (46.0/50.0) G2: 50.2 (48.6/51.9) G3: 49.0 (46.5/51.4) Physical component summary: G1: 43.2 (39.4/47.1) G2: 45.1 (42.7/47.4) G3: 44.0 (41.8/46.1) General health profile: G1: 63.4 (53.8/68.5) G2: 66.8 (63.5/70.2) G3: 62.0 (57.9/66.0)	No significant differences in scale scores were seen across the arms at baseline or after the intervention Mental summary: G1 vs. G2: 2.0* G1 vs. G3: 2.1* G2 vs. G3: 0.1* p=NR Physical summary G1 vs. G2: 1.4* G1 vs. G3: 0.8* G2 vs. G3: 2.2* p=NR General health: G1 vs. G2: 3.9* G1 vs. G3: 1.6* G2 vs. G3: 5.5* p=NR 12 month follow-up Mental summary: G1 vs. G2: 2.2* G1 vs. G3: 1.0* G2 vs. G3: 1.2* p=NR Physical summary G1 vs. G2: 1.9* G1 vs. G3: 0.8* G2 vs. G3: 1.1* p=NR General health: G1 vs. G2: 3.4* G1 vs. G3: 1.4* G2 vs. G3: 4.8* p=NR	deprivation, age, sex, and the training status of the practice t tests

Table E-5. Key question 2 studies first outcome (continued)

Author, Year	Groups	Outcome #1, Exact Measure Used	Timing of Measurement, Data Source	N analyzed for This Outcome	Results by Group	Differences in Groups	Covariates Controlled for in Analysis, Statistical Methods Used
Eaton et al., 2011[11]	G1: 1-hour academic detailing (increase clinician ability) G2: Academic detailing plus a patient education toolkit, a computer kiosk with patient activation software, and a PDA-based decision support tool (multi-component)	Clinical outcomes (applicable for general public/patients) Percentage of patients screened for hyperlipidemia and treated to their LDL and non–HDL cholesterol goals	Baseline and one year post-intervention Objective measurement (medical records) and self-report (by physicians)	4,105 patients G1: 2,000 G2: 2,105	Both groups improved screening (89%) and the percentage of patients at their LDL (74%) and non-HDL cholesterol goals (74%), p < .001. Results by group, p=NR	No significant difference between groups for primary outcome. Post hoc analysis: G2: Difference: Practices with above-median use of the patient activation kiosk were more likely to have patients screened with a full lipid profile OR: 2.54 95% CI: 1.97 to 3.27 p=NR Difference: Physicians who were more frequent users of the PDA decision support tool were more likely to have their patients at LDL cholesterol goals (16%) OR = 1.16 95% CI: 0.98 to 1.36 Difference: Physicians who were more frequent users of the PDA decision support tool were more likely to have their patients at LDL cholesterol	None Generalized linear mixed model

Table E-5. Key question 2 studies first outcome (continued)

Author, Year	Groups	Outcome #1, Exact Measure Used	Timing of Measurement, Data Source	N analyzed for This Outcome	Results by Group	Differences in Groups	Covariates Controlled for in Analysis, Statistical Methods Used
Eaton et al., 2011[11] (continued)						goals (16%) OR,1.27; 95% CI, 1.07-1.50 and non-HDL cholesterol goals (12%) OR = 1.12 95% CI: 0.95-1.32	
Elder et al., 2005[12]	G1: Culturally targeted print-materials + activity inserts (increase reach) G2: Tailored print materials + activity inserts + supporting materials (multi-component). G3: Tailored print materials + in-person promotora (multi-component)	Clinical outcomes (applicable for general public/patients) % calories from fat	Baseline and 12 week follow-up Self-report face-to-face interview	Baseline N = 357 G1: 119 G2: 118 G3: 120 Follow-up N = 313 G1: 107 G2: 99 G3: 107	Adjusted Mean at Time 2 % calories from fat: G1: 30% G2: 30.4% G3: 29.3%	G1 vs. G2: 0.4%* G2 vs. G3: 1.1%* G1 vs. G3: 0.7%* p=NR, but it was not significant.	Baseline measure Tukey-Kramer multiple comparison test

Table E-5. Key question 2 studies first outcome (continued)

Author, Year	Groups	Outcome #1, Exact Measure Used	Timing of Measurement, Data Source	N analyzed for This Outcome	Results by Group	Differences in Groups	Covariates Controlled for in Analysis, Statistical Methods Used
Feldstein et al., 2006[13]	G1: Usual care (not abstracted) G2: EMR reminder (increase reach for clinicians) G3: EMR reminder and patient reminder (via letter with educational materials (multi-component)	Health-related decisions or behavior (applicable for general public/patients) Percent receiving pharmacological treatment defined as drugs dispensed to patient from outpatient pharmacy system	Within 6 months of intervention Objective measure from pharmacy system	G1: 101 G2: 101 G3: 109	G1: 4.0% G2: 11.9% G3: 10.1%	Difference: G2 vs. G1 7.9 95% CI: .47 (.35-.59) p=NR G3 vs. G1 6.1 95% CI: .38 (.26-.50) p=NR G3 vs. G2 -2.2 95% CI:NR	Fracture type, age, weight less than 127 pounds, osteoporosis diagnosis, and Charlson co-morbidity index. General linear modeling using treatment group, fracture type, age, weight, osteoporosis diagnosis and Charlson Comorbidity Index indicators. Models include independent variables significant in univariate analyses at p<.10. Continuous outcome measures change scores regressed on the baseline values and indicators of treatment groups. Logistic regression used for unadjusted results.

Table E-5. Key question 2 studies first outcome (continued)

Author, Year	Groups	Outcome #1, Exact Measure Used	Timing of Measurement, Data Source	N analyzed for This Outcome	Results by Group	Differences in Groups	Covariates Controlled for in Analysis, Statistical Methods Used
Gattellari et al., 2005[14]	G1: Leaflet (increase reach) G2: Video (increase reach) G3: Booklet (increase reach)	Knowledge about the evidence 14-item measure comprised of 10 T/F questions and 4 multiple choice questions administered at pre and post test - 2 items on efficacy of PSA screening; 3 on test accuracy; 1 on controversy about PSA screening; 4 on nature of prostate cancer; 2 on risk factors for prostate cancer, and 2 on treatment-related issues; scores were summed and multiplied by 100 for % of items correctly answered	Mean 21 days after receiving information (range 15 to 31) Self-report	N=405	Pre-test: G1: 30.1% G2: 28.7% G3: 29.8% Post-test: G1: 42.2% G2: 45.8% G3: 57.2%	Absolute differences within arms (pre-post): G1: 12.1%*, CI and p <0.001 G2: 17.1%*, CI and p<0.001 G3: 27.9%*, CI and p<0.001 Absolute difference in changes between arms: G2-G1: 5.0%*, CI and p=NR G3-G1: 15.8%*, CI and p=NR G3-G2: 10.8%*, CI and p=NR Post-test G2-G1: 3.6%* Post-test G3-G1: 15.0%* Post-test G3-G2: 11.4%* Overall p<0.001	None Wilcoxon signed rank test and ANOVA

Table E-5. Key question 2 studies first outcome (continued)

Author, Year	Groups	Outcome #1, Exact Measure Used	Timing of Measurement, Data Source	N analyzed for This Outcome	Results by Group	Differences in Groups	Covariates Controlled for in Analysis, Statistical Methods Used
Hagmolen et al., 2008[15]	G1: Guideline dissemination (increase reach) G2: Guideline dissemination + educational program (increase ability) G3: Guideline dissemination + educational program + individualized treatment advice based on airway responsiveness and symptoms (multi-component)	Clinical outcomes (applicable for general public/patients) Change in AHR: reflects severity of the asthma. A single concentration methacholine challenge test was performed when FEV% predicted was greater than or equal to 75%. The degree of AHR was expressed as a PD20. Moderate to severe AHR was defined as a PD20 of less than or equal to 300 mcg.	Baseline and one year follow-up; one year between measures Objective measurement	Overall N: 362 G1: 98 G2: 133 G3: 131 Also conducted post-hoc analysis where Groups 1 and 2 were combined	G1: M = 8.3 (SE = 0.2) G2: M = 8.2 (SE = 0.2) G3: M = 8.7 (SE = 0.2) Post-hoc analysis: G1&G2: M = 8.3 (SE=0.2) G3: M = 8.7 (SE=0.2)	Difference: G1 vs. G2: 0.1* G1 vs. G3: 0.4* G2 vs. G3: 0.5* No significant differences between all 3 groups p=0.09 Significant difference between baseline and end of study for G3: 0.7, p=0.001 Post-hoc analysis (aggregated groups 1 & 2): G1&G2 vs. G3: 0.4* Significant difference between groups p=0.03 Significant difference between baseline and end of study for G1&G2 combined: 0.27, p=0.05 and G3: .7, p<0.001.	NR Mixed model ANOVA analyses

Table E-5. Key question 2 studies first outcome (continued)

Author, Year	Groups	Outcome #1, Exact Measure Used	Timing of Measurement, Data Source	N analyzed for This Outcome	Results by Group	Differences in Groups	Covariates Controlled for in Analysis, Statistical Methods Used
Jain et al., 2006[16]	G1: Passive intervention-guidelines by mail (increase reach) G2: Active intervention (multi-component)	Behavior (applicable for clinicians) Nutritional adequacy of EN. Defined as the calories received from EN divided by the maximum total daily calories prescribed (recommended by the dietitian) for each individual patient during the first 12 days of ICU stay.	Baseline, 12 month follow-up Chart	Practice Overall: 58 ICUs randomized as 50 clusters G1: 25 clusters G2: 25 clusters Patients Baseline Overall: 623 G1: 298 G2: 325 Follow-up Overall: 612 G1: 305 G2: 307 Note: the patients were not the same at baseline and follow-up. The authors took a cross-sectional survey at both time points.	Baseline Mean ± SE G1: 45.2 ± 2.5 G2: 40.7 ± 2.5 Follow-up G1: 51.3 ± 2.6 (change from baseline: 6.2 ± 2.2, P =0.005) G2: 48.7 ± 2.6 (change from baseline 8.0 ± 2.1, P<0.001)	Baseline Difference (G1- G2) Mean ± SE - 4.5 ± 3.5 Follow-up Difference (G1- G2) Mean ± SE - 2.6 ± 3.5 Change 1.9 ± 3.1, p=0.541 In Subgroup analysis of medical patients only, the difference was significant. Difference in change from baseline to follow-up between groups: 8.1 ± 3.9, p=0.036	ICU length of stay Two-level hierarchical model as implemented in HLM

Table E-5. Key question 2 studies first outcome (continued)

Author, Year	Groups	Outcome #1, Exact Measure Used	Timing of Measurement, Data Source	N analyzed for This Outcome	Results by Group	Differences in Groups	Covariates Controlled for in Analysis, Statistical Methods Used
Jousimaa et al., 2002[17]	G1: Computerized version of guidelines (increase ability) G2: Textbook-based version of guidelines (increase reach)	Behavior (applicable for clinicians) Number (and percent) of relevant consultations compliant with guidelines	One month post-intervention Objective and self-report	Laboratory examinations: Overall N: G1: 1640 G2: 1529 Radiological examinations: Overall N: G1: 1604 G2: 1518 Physical examinations: Overall N: G1: 1610 G2: 1545 Other examinations: Overall N: G1: 314 G2: 307 Procedures: Overall N: G1: 196 G2: 171 Nonpharma-cologic treatment: Overall N: G1: 92 G2: 122 Pharmacologic treatments: Overall N: G1: 1654 G2: 1568	Laboratory examinations: G1: 1481 (90.3%) G2: 1372 (89.7%) Radiological examinations: G1: 1504 (93.8%) G2: 1416 (93.3%) Physical examinations: G1: 1494 (92.8%) G2: 1461 (94.6%) Other examinations: G1: 235 (74.8%) G2: 248 (80.8%) Procedures: G1: 152 (77.6%) G2: 140 (81.9%) Nonpharmacologic treatment: G1: 80 (87.0%) G2: 110 (90.2%) Pharmacologic treatments: G1: 1391 (84.1%) G2: 1350 (86.1%) Physiotherapy: G1: 77 (78.6%) G2: 83 (80.6%) Referrals: G1: 1619 (96.1%) G2: 1508 (95.6%)	Proportion of noncompliant decisions considered to be clinically important (major or serious) similar in the two groups: 47.4% (407/859) in G1 compared with 46.3% (349/753) G2. No statistically significant differences between groups in terms of compliance. Laboratory exams: G1 vs. G2: 109 Difference: OR = 1.07 95% CI: 0.79-1.44 ICC: 0.015 Radiological exams: G1 vs. G2: 88 Difference: OR = 1.09 95% CI: 0.81-1.46 ICC: 0 Physical examinations: G1 vs. G2: 33 Difference: OR = 0.74 95% CI: 0.51-1.06 ICC: 0.015	NR Chi-squared tests; a retrospective power calculation, adjusting for clustering using an ICC of 0.015 and an average cluster size of 27

Table E-5. Key question 2 studies first outcome (continued)

Author, Year	Groups	Outcome #1, Exact Measure Used	Timing of Measurement, Data Source	N analyzed for This Outcome	Results by Group	Differences in Groups	Covariates Controlled for in Analysis, Statistical Methods Used
Jousimaa et al., 2002[17] (continued)				Physio-therapy: Overall N: G1: 98 G2: 103 Referrals: Overall N: G1: 1684 G2: 1578		Other examinations: G1 vs. G2: 13 Difference: OR = 0.71 95% CI: 0.43-1.17 ICC: 0.021 Procedures: G1 vs. G2: 12 Difference: OR = 0.77 95% CI: 0.43-1.36 ICC: 0 Nonpharmacologic treatment: G1 vs. G2: 30 Difference: OR = 0.73 95% CI: 0.22-2.41 ICC: 0.058 Pharmacologic treatments: G1 vs. G2: 41 Difference: OR = 0.85 95% CI: 0.67-1.09 ICC: 0.010 Physiotherapy: G1 vs. G2: 6 Difference: OR = 0.88 95% CI: 0.34-2.32 ICC: 0.195 Referrals: G1 vs. G2: 111 Difference: OR = 1.13 95% CI: 0.79-1.63 ICC: 0.002	

Table E-5. Key question 2 studies first outcome (continued)

Author, Year	Groups	Outcome #1, Exact Measure Used	Timing of Measurement, Data Source	N analyzed for This Outcome	Results by Group	Differences in Groups	Covariates Controlled for in Analysis, Statistical Methods Used
Junghans et al., 2007[18]	G1: Conventional guideline (increase reach) G2: Ratings about specific patients in vignettes (increase motivation)	Behavior (applicable for clinicians) Agreement of physicians' recommendations with those made by 2 independent expert panels. Agreement was defined by a physician recommending definitely or probably doing a test rated appropriate by the panels or by recommending definitely or probably not doing a test rated inappropriate. An unsure recommendation was interpreted as disagreement	Baseline and immediate posttest Self-reported decision	N = 292 G1: 147 G2: 145	% that had an appropriate Baseline Exercise ECG G1: 42.7% G2: 43.5% Angiography G1: 64.9% G2: 64% Post-intervention Exercise ECG decision G1: 43.5% G2: 54.9% Angiography G1: 64% G2: 79.9%	Between-arm comparisons Odds Ratio (95%CI), P value Patient-specific ratings Exercise ECG OR: 1.57 (1.36,1.82), P<0.001 Angiography OR: 2.24 (1.90,2.62), P<0.001 Convential guidelines Exercise ECG OR: 0.96 (0.83,1.11), P<0.001 Angiography OR: 1.05 (0.87,1.26), P<0.001	NR random-effects logistic regression analysis allowing for intracluster correlation

Table E-5. Key question 2 studies first outcome (continued)

Author, Year	Groups	Outcome #1, Exact Measure Used	Timing of Measurement, Data Source	N analyzed for This Outcome	Results by Group	Differences in Groups	Covariates Controlled for in Analysis, Statistical Methods Used
Kennedy et al., 2003[19]	G1: Control (not abstracted) G2: Information (increase reach) G3: Interview (increase motivation)	Clinical outcomes (applicable for general public/patients) Health Status. Measured using the 36-item short-form general health survey (SF-36) instrument	6 and 12 month data merged together to for a "short-term" follow-up dataset. 24 months is labeled "long-term" Self-report	Overall: 595 G2: 198 (97% completed) G3: 208 (94% completed)	NR	Adjusted mean between-group difference (G2 vs. G3) at short-term follow-up (95% CI) Physical function: 0.0 (-3.5/ 3.5) Social function: -2.7 (-7.7/2.2) Role physical: -2.5 (-10.3/5.2) Role emotional: -4.6 (-13.9/13.7) Mental health: -2.5 (-6.6/1.6) Energy: -2.5 (-6.9/2.0) Pain: -1.3 (-6.4/3.9) General health perception: -0.8 (-5.2/3.5) Adjusted mean between-group difference (G2 vs. G3) at long-term follow-up (95% CI) PF: -1.5 (-5.2/2.3) SF: 3.2 (-1.6/8.1) RP: 5.7 (-2.1/13.6) RE: 7.1 (-2.0/16.4) MH: 1.1 (-2.8/4.9) Energy: 0.4 (-5.0/5.7) Pain: 0.3(-5.2/5.7) GHP: -0.1 (-4.0/3.7)	Consultant sex; Consultant year of qualification; Age; Baseline health status score; Baseline menorrhagia severity; Baseline knowledge; Duration of problem; Length of follow-up Multiple regression

Table E-5. Key question 2 studies first outcome (continued)

Author, Year	Groups	Outcome #1, Exact Measure Used	Timing of Measurement, Data Source	N analyzed for This Outcome	Results by Group	Differences in Groups	Covariates Controlled for in Analysis, Statistical Methods Used
King et al., 2007[20]	G1: Attention control (not abstracted) G2: Counselor via phone (increase motivation) G3: Automated counselor via phone (increase reach)	Clinical outcomes (applicable for general public/patients) Physical activity behavior. Assessed using the Stanford 7-Day PAR. The PAR-based mean daily energy expenditure estimates from MOD activity was the primary study outcome measure (#1 below). Measures: (1) PAR energy expenditures in moderate-intensity or more vigorous (MOD+) activity, kcal/kg-1/day-1 (SD) (2) PAR minutes of MOD+ activity/wk, Mean (SD) (3) PAR days/wk engaged in ≥ 30 min of MOD+ activity, Mean(SD)	Baseline, 6, 12 months Self report	N = 189 G2: 66 G3: 61	PAR kcal/kg-1/day-1 (SD) Baseline G2: 0.85 (1.0) G3: 0.80(1.2) 6 months G2: 1.69 (1.1) G3: 1.53 (1.3) 12 months G2: 1.64 (1.3) G3: 1.56 (1.4) PAR min. of MOD+ activity/wk Baseline G2: 99.7 (147.6) G3: 78.4 (113.3) 6 months G2: 170.7 (104.4) G3: 180.0 (230.6) 12 months G2: 177.8 (133.6) G3: 157.3 (142.9) PAR days/wk engaged in ≥ 30 min of MOD+ Baseline G2: 1.4(1.5) G3: 1.1 (1.6) 6 months G2: 3.2 (2.0) G3: 2.6 (2.3) 12 months G2: 3.1 (2.0) G3: 2.8 (2.5)	Changes across 6 months PAR kcal/kg-1/day-1 G2 vs. G3: 0.11* p=0.73 PAR min. of MOD+ activity/week G2 vs. G3: 9.3*, p=0.65 PAR days/week engaged in ≥ 30 min of MOD+ G2 vs. G3: 0.3*, p=NR but it was ns Changes across 12 months kcal/kg-1/day-1 (SD) G2 vs. G3=0.08*, p=0.60 PAR min. of MOD+ activity/week G2 vs. G3: 20.5*, p=0.66 PAR days/week engaged in ≥ 30 min of MOD+ G2 vs. G3: 0.3*, p=NR but it was ns	Baseline levels of dependent variables Gender ANCOVA

Table E-5. Key question 2 studies first outcome (continued)

Author, Year	Groups	Outcome #1, Exact Measure Used	Timing of Measurement, Data Source	N analyzed for This Outcome	Results by Group	Differences in Groups	Covariates Controlled for in Analysis, Statistical Methods Used
Laprise et al., 2009[21]	G1: CME (increase ability) G2: CME + practice enablers and reinforcers (multi-component)	Behavior (applicable for clinicians) Adherence to guidelines. Proportion of patients, undermanaged at baseline for at least 1 recommendation, for which study physicians undertook at least 1 preventive-care action in the first visit following patients' recruitment in the study. A binary outcome was used.	Baseline and followup (exact time not specified) Retrospective audit information	G1: 948 G2: 1396	Baseline # of undermanaged rec/patient, n (%) None G1: 172 (18.1%) G2: 263 (18.8%) 1 G1: 313 (33.0%) G2: 452 (32.4%) 2 G1: 282 (29.7%) G2: 339 (32.2%) 3-5 G1: 181 (19.1%) G2: 232 (16.6%) Follow-up Implementation of at least 1 of the secondary outcomes G1: 225 (29.0%) G2: 474(41.8%)	Odds Ratio (95% CI) 1.78 (1.32-2.41)	NR Logistic regression

Table E-5. Key question 2 studies first outcome (continued)

Author, Year	Groups	Outcome #1, Exact Measure Used	Timing of Measurement, Data Source	N analyzed for This Outcome	Results by Group	Differences in Groups	Covariates Controlled for in Analysis, Statistical Methods Used
Lien et al., 2007,[22] Svetkey et al., 2003,[23] Young et al., 2009[24]	G1: Advice only (increase reach) G2: Advice + behavioral counseling using established intervention (multi-component) G3: Established intervention + DASH dietary recommendations (multi-component)	Clinical outcomes (applicable for general public/patients) Change in BP. SBP was the appearance of the first Korotkoff sound; DBP was the disappearance of Korotkoff sounds. At each assessment point, BP was the mean of all of the available measurements. Per criteria: good levels of BP are ≥130/ ≥85 mm Hg	Baseline and at 6-month follow-up. 2 BP measurements separated by 30 seconds were obtained and averaged/ Objective measurement	Overall N: 671 G1: 273 G2: 188 (71% of randomized participants) G3: 210 (78% of randomized participants)	Reduction from baseline to 6 month follow-up for SBP G1: 6.6 (9.2) mm Hg G2: 10.5 (10.1) mm Hg G3: 11.1 (9.9) mm Hg Reduction from baseline to 6 month follow-up for DBP G1: 3.8 (6.3) mm Hg G2: 5.5 (6.7) mm Hg G3: 6.4 (6.8) mm Hg	On Treatment Analysis Change (Δ) in BP between-group differences (Mean and CI) Δ in G2 minus Δ in G1: -4.9 (-6.6 to -3.3) P<0.001 Δ in G3 minus Δ in G1: -5.7 (-7.2 to -4.1) p<0.001 Δ in G3 minus Δ in G2: -0.7 (-2.5 to 1.0) p=0 .41 Change(Δ) in Diastolic BP between-group differences (Mean and CI) ΔG2-ΔG1: -2.5 (-3.7 to -1.3), p<.001 ΔG3 - ΔG1:: -3.2 (-4.3 to -2.0), p<.001 ΔG3-ΔG2: -0.7 (-1.9 to 0.6), p=0.29	Age, gender, race General linear modeling

Table E-5. Key question 2 studies first outcome (continued)

Author, Year	Groups	Outcome #1, Exact Measure Used	Timing of Measurement, Data Source	N analyzed for This Outcome	Results by Group	Differences in Groups	Covariates Controlled for in Analysis, Statistical Methods Used
Lien et al., 2007,[22] Svetkey et al., 2003,[23] Young et al., 2009[24] (continued)						Intention to Treat Analysis SBP ΔG2-ΔG1: -3.7 (-5.3 to -2.1), P<0.001 ΔG3-ΔG1: -4.3 (-5.9 to -2.8) P<0.001 ΔG3-ΔG2: -0.6 (-2.2 to 0.9) p=0.43 DBP ΔG2-ΔG1: -1.7 (-2.8 to -.06), P<0.01 ΔG3-ΔG1: -2.6 (-3.7 to -1.5), P<0.001 ΔG3-ΔG2: -0.9 (-2.0 to 0.2), p=0.11	

Table E-5. Key question 2 studies first outcome (continued)

Author, Year	Groups	Outcome #1, Exact Measure Used	Timing of Measurement, Data Source	N analyzed for This Outcome	Results by Group	Differences in Groups	Covariates Controlled for in Analysis, Statistical Methods Used
Marcus et al., 2009[25]	G1: Contact control treatment delayed group (not abstracted) G2: Telephone-based individualized feedback (increase motivation) G3: Print-based individualized feedback (increase reach)	Behavioral intentions to use or apply the evidence Instrument developed for behavioral processes of change for exercise by Marcus, et al.	Baseline, 6 and 12 months Self-report	NR	G1: 6 Months: 2.43; 12 Months: 2.41 G2: 6 Months: 3.08; 12 Months: 2.82 G3: 6 Months: 2.95; 12 Months: 2.91	Difference: 6 Months: F=24.01; 12 Months: 13.73 95% CI: NR 6 Months: P<0.0001 12 Months: P<0.0001	Yes Analysis of covariance, adjusted for treatment effects for gender and seasonal differences. When overall test of between-groups differences was significant at the >05 level, the source of these differences was examined further using single-degree-of-freedom contrasts that compared the active treatment arms with each other as well as with the treatment delayed group.

Table E-5. Key question 2 studies first outcome (continued)

Author, Year	Groups	Outcome #1, Exact Measure Used	Timing of Measurement, Data Source	N analyzed for This Outcome	Results by Group	Differences in Groups	Covariates Controlled for in Analysis, Statistical Methods Used
Maxwell et al.,2010[26]	G1: Control (not abstracted) G2: Educational session + letter to provider (multi-component) G3: Educational session + letter to provider + FOBT kit (multi-component)	Clinical outcomes (applicable for general public/patients) Self-reported screening NOTE: participants w/out outcome data were classified as not-screened	6 months Self-report NOTE: subsample validated by physician report for 141 patients	542, but imputed information on 110 of them (20%)	G1: 14 (9%) G2: 45 (25%) G3: 61 (30%)	G2 v. G3 Difference: 5% 95% CI: NR p= NR OR G2 to G1 (95% CI): 3.7 (1.8, 7.5) P<0.001 OR G3 to G1 (95% CI): 4.9 (2.4, 9.9) P<0.001	adjusted for baseline imbalance (e.g. language of baseline interview) and clustering within organization and session mixed effects model w/random intercepts for organizations and session within organization

Table E-5. Key question 2 studies first outcome (continued)

Author, Year	Groups	Outcome #1, Exact Measure Used	Timing of Measurement, Data Source	N analyzed for This Outcome	Results by Group	Differences in Groups	Covariates Controlled for in Analysis, Statistical Methods Used
Murtaugh et al.,2005[27]	G1: Usual care (not abstracted) G2: Basic intervention e-mail reminder (increase reach) G3: Augmented intervention of e-mail reminder + package of supporting materials (multi-component)	Discussions about the evidence % giving patients global instructions about signs and symptoms of CHF	Chart-review of subsequent RN visit, within 45 days of initial intake chart	354	Overall N: 354 G1: 42.1% G2: 53.9% G3: 59.5%	Difference G2-G1: 11.8%, p=0.070 Difference G3-G1: 17.4%, p=0.007 Difference G3-G2: 5.6%*, CI and p=NR	Sociodemo-graphic variables of the RN (age, gender, race/ethnicity), Rn employment status, educational level and caseload; average baseline characteristics of patients care for by each RN including health, functional status; geographic area where nurse provided care Predictive multivariate modeling

Table E-5. Key question 2 studies first outcome (continued)

Author, Year	Groups	Outcome #1, Exact Measure Used	Timing of Measurement, Data Source	N analyzed for This Outcome	Results by Group	Differences in Groups	Covariates Controlled for in Analysis, Statistical Methods Used
Paradis et al.,2011[28]	G1: Paper handouts (increase reach) G2: Educational DVD (increase reach)	Knowledge about the evidence Knowledge of infant development; measured using a subset of 14 questions from the 58-item Knowledge of Infant Development Inventory that pertained most to newborns. Answers were scored as correct or incorrect. Parents could answer each statement with “agree,” “disagree,” or “not sure,” with uncertain answers considered incorrect.	2 weeks post-intervention Self-report	Overall N: 137 G1: 67 G2: 70	Mean change in Knowledge (from baseline): G1: -0.06 (SD =2.99) G2: 0.00 (SD=2.53) NOTE: baseline scores G1: 10.2 G2: 9.4	G2-G1: -0.06 p=0.90	Hispanic ethnicity, babies born at outside hospital, #exclusively breast fed multivariate regression analysis

Table E-5. Key question 2 studies first outcome (continued)

Author, Year	Groups	Outcome #1, Exact Measure Used	Timing of Measurement, Data Source	N analyzed for This Outcome	Results by Group	Differences in Groups	Covariates Controlled for in Analysis, Statistical Methods Used
Partin et al., 2004[29]	G1: Usual care (not abstracted) G2: Pamphlet (increase reach) G3: Video (increase reach)	Knowledge about the evidence CaP screening knowledge, as assessed from a 10-item index. The index score is calculated as the summative number of correct responses to 10 knowledge questions. "Don't know" responses are treated as incorrect. Index scores range from 0 to 10	1 week post- target appointment Self-report	N = 893 G2: 295 G3: 308	CaP knowledge index: mean scores: G2: 7.3 G3: 7.4 Other CaP screening knowledge items (Unadjusted) PSA predictive value G2: 0.22 G3: 0.28 Natural History G2: 0.61 G3: 0.62 Treatment efficacy G2: 0.20 G3: 0.19 Expert disagreement G2: 0.18 G3: 0.29	CaP Index: G2 vs. G3: 0.1*, p=NR Other CaP knowledge items: PSA predictive value G2 vs. G3: 0.06*, ns Natural History G2 vs. G3: 0.01*, ns Treatment efficacy G2 vs. G3: 0.01*, ns Expert disagreement G2 vs. G3: .11*, p=0.009	Baseline characteristics Logistic regression and standard linear regression

Table E-5. Key question 2 studies first outcome (continued)

Author, Year	Groups	Outcome #1, Exact Measure Used	Timing of Measurement, Data Source	N analyzed for This Outcome	Results by Group	Differences in Groups	Covariates Controlled for in Analysis, Statistical Methods Used
Rahme et al., 2005[30]	G1: No treatment control (not abstracted) G2: Decision tree (increase ability) G3: Workshop (increase ability) G4: Workshop + decision tree (multi-component)	Behavior (applicable for clinicians) Retrospective assessment of prescribing. A score of zero or 1 was given to every prescription that was judged as adequate according to the decision tree.	5-months prior to intervention/ 5-months post intervention Objective measurement: Data were obtained from the Provincial Health Care Fund database	N of prescriptions Pre-intervention: Total: 5318 G2: 1569 G3:536 G4: 1776 Post-intervention Total: 4610 G2: 1317 G3: 450 G4: 1634	Pre-intervention G2: 51% G3: 51% G4: 58% Post-intervention G2: 54% G3: 56% G4: 62%	Only compared groups to control: Ratio of OR (95%CI) G2 vs. CRL: 1.0 (0.6/1.7) G3 vs. CTRL: 5.7 (0.4/26.9) G4 vs. Ctrl: 1.9 (0.9/3.8) Within-group differences (post vs. pre) G2: 1.3 (0.9-1.8) G3: 1.6 (0.9-1.8) G4: 1.8 (1.3-2.4)	Risk of gastrointestinal even. Additional analyses: Per protocol analysis excluding physicians in the workshop and workshop and tree group who did not attend the workshop Multi-level Bayesian hierarchical model

Table E-5. Key question 2 studies first outcome (continued)

Author, Year	Groups	Outcome #1, Exact Measure Used	Timing of Measurement, Data Source	N analyzed for This Outcome	Results by Group	Differences in Groups	Covariates Controlled for in Analysis, Statistical Methods Used
Rebbeck et al., 2006[31]	G1: Dissemination of guidelines by mail (increase reach) G2: Implementation group (multi-component)	Clinical outcomes (applicable for general public/patients) Disability - measured using the Functional Rating Index which measures disability due to back and neck pain. It is a 10-item questionnaire with a 5-point response scale for each item. Summation of the 10 items yields a score ranging from 0 to 40, with higher scores indicating greater perceived disability.	Baseline, month 1.5, month 3, month 6, month 12 Self-report	Baseline: G1: 28 G2: 71 Month 1.5 G1: 24 G2: 64 Month 3 G1: 23 G2: 59 Month 6 G1: 19 G2: 56 Month 12 G1: 26 G2: 67	Baseline: G1: M=23.9, SD=8.6 G2: 22.8, SD=8.2 Month 1.5 G1: 14.8, SD=8.8 G2: 15.8, SD=8.7 Month 3 G1: 12.8, SD=8.5 G2: 12.7, SD=8.5 Month 6 G1: 11.3, SD=9.3 G2: 11.5, SD=9.0 Month 12 G1: 12.0, SD=10.4 G2: 11.4, SD=8.9	Baseline Difference (G1 vs. G2): 1.0* 95% CI: -6.1 to 4.1 p=0.68 Month 1.5 Difference (G1 vs. G2): 1.0* 95% CI: -5.1 to 7.1 p=0.74 Month 3 Difference (G1 vs. G2): 0.1* 95% CI: -5.8 to 5.7 p=0.99 Month 6 Difference (G1 vs. G2): 0.1* 95% CI: -6.4 to 6.7 p=0.97 Month 12 Difference (G1 vs. G2): 0.6* 95% CI: -7.8 to 6.6 p=0.87	NR T-test, adjusted using methods for cluster-randomized trials

Table E-5. Key question 2 studies first outcome (continued)

Author, Year	Groups	Outcome #1, Exact Measure Used	Timing of Measurement, Data Source	N analyzed for This Outcome	Results by Group	Differences in Groups	Covariates Controlled for in Analysis, Statistical Methods Used
Rimer et al., 2001[32]	G1: No treatment control/usual care (not abstracted) G2: Tailored print (increase reach) G3: Tailored print + telephone counselling (multi-component)	Health-related decisions or behavior (applicable for general public/patients) Receipt of a mammogram yearly	Interview 15 months after receiving intervention Self-report	Overall N: 1127 G1: 412 G2: 392 G3: 323	Baseline- percent up-to-date NR Follow-up mammogram in 15 months: G1: 260*, 63% G2: 239*, 61% G3: 223*, 69%	Overall p=0.066 G2-G1:- 2%*, NS G3-G1: 6%*, NS G3-G2: 8%*, NS	None Pearson chi-squared; F-test

Table E-5. Key question 2 studies first outcome (continued)

Author, Year	Groups	Outcome #1, Exact Measure Used	Timing of Measurement, Data Source	N analyzed for This Outcome	Results by Group	Differences in Groups	Covariates Controlled for in Analysis, Statistical Methods Used
Simon et al., 2005[33]	G1: Mailed educational materials (increase reach) G2: Individual academic detailing (increase ability) G3: Group academic detailing (increase ability)	Behavior (applicable for clinicians) Change in guideline adherence - A patient was considered to have received a diuretic or beta blocker if he or she received at least one prescription for either drug during the specified time frame.	Baseline, 1-year follow-up, 2-year follow-up Objective measurement (prescription via claims)	Baseline: 3692 Year 1: 3556 Year 2: 2572	Percent increase Year 1 G1: 6.2% G2: 12.5% G3: 13.2% Year 2 G1: 10.1% G2: 14.7% G3: 11.3%	Year 1 G1 vs. G3: 7%* Difference: Diuretic or beta blocker use was more likely in G3 than G1 (OR, 1.40) 95% CI: 1.11-1.76 p= NR G1 vs. G2: 6%* Difference: Diuretic or beta blocker use was more likely in G2 than G1 (OR, 1.30) 95% CI: 0.95-1.79 P:=NR Year 2 G1 vs. G2: 4.6% Difference: Diuretic or beta blocker use was more likely in G2 than G1 (OR, 1.22) 95% CI: 0.92-1.62 p=ns G1 vs. G3: 1.2% Difference: Diuretic or beta blocker use was not more likely in G3 than G1 (OR, 1.06) 95% CI: 0.80-1.39 p=NR	Differences among individual patients Logistic regression

Table E-5. Key question 2 studies first outcome (continued)

Author, Year	Groups	Outcome #1, Exact Measure Used	Timing of Measurement, Data Source	N analyzed for This Outcome	Results by Group	Differences in Groups	Covariates Controlled for in Analysis, Statistical Methods Used
Soler et al., 2010[34]	G1: Control (not abstracted) G2: Training session on the SEPAR guidelines (increase ability) G3: G2 + portable-device for spirometry (multi-component)	Clinical outcomes (applicable for general public/patients) Changes in COPD stratification and diagnostic testing according to SEPAR guidelines	Adequate COPD classification according to SEPAR guidelines Chart	G1: 1481, G2: 2119, G3: 5556 (Phase II)	G1: 60.1% G2: 69% G3: 88.5%	Absolute difference in accurate stratification: G2-G1: 8.9%, p=NR G3-G1: 28.4%, p=NR	Baseline variable Within group changes in the three groups assessed by ANCOVA; b/t group p-values NR
Sullivan et al., 2010[35]	G1: VA guidelines (increase reach) G2: COPE: web-based education program (increase ability)	Knowledge about the evidence Knowledge of the role of opioids in CNCP was assessed with 9 multiple choice board-style questions developed by the authors covering opioid pharmacology, controlled substance regulations, and diagnostic challenges (range 0-9)	Pre-training and immediately post-training Self-report	N = 159	G1: Pretest: M = 5.7, SD = 1.3 Posttest: M = 6.1, SD = 1.3 G2: Pretest: M = 5.9, SD = 1.4 Posttest: M = 8.4, SD = 0.8	G1 vs. G2 (posttest): 2.3* Difference: t = 12.41, P <0.001 Difference: Significant time by group interaction (different rates of change over time) (Wald $\chi 2$ = 72.06, df = 1, P<0.00001)	Gender; year of residency (no effects observed for these variables) Independent group t tests; intention-to-treat analyses using the GEE

Table E-5. Key question 2 studies first outcome (continued)

Author, Year	Groups	Outcome #1, Exact Measure Used	Timing of Measurement, Data Source	N analyzed for This Outcome	Results by Group	Differences in Groups	Covariates Controlled for in Analysis, Statistical Methods Used
Watson et al., 2002[36]	G1: Guideline materials by postal mail (increase reach) G2: EO session and guidelines (increase ability) G3: CPE session and guidelines (increase ability) G4: Guidelines + EO and CPE (multi-component)	Behavior (applicable for clinicians) Appropriateness of OTC management of vulvovaginal candidiasis by community pharmacy staff: measured by the proportion of visits resulting in an appropriate sale or non-sale of an anti-fungal product (based upon the guideline recommendations)	Ten local amateur actors conducted simulated patient visits with 7 scenarios. Each pharmacy was visited 7 times; twice before the intervention between March and April 2000 and five times after the intervention between July and November 2000. No pharmacy received more than one visit per month. Following each visit, the actor completed an assessment form, recording details of their visit, including sale/no sale, product details and the number of staff involved in the interaction. Direct observation and assessment	Baseline: G1: 27 visits; G2: 27 visits; G3: 27 visits; G4: 27 visits Follow-up: G1: 69 visits G2:69 visits G3: 69 visits G4: 69 visits	Baseline: Appropriate Outcome: G1: 10 (37%); G2: 11 (41%); G3: 10 (37%) G4: 10 (37%) Followup: Appropriate Outcome: G1: 24 (35%); G2: 32 (46%); G3: 25 (36%); G4: 24 (35%)	Difference: G2 EO vs. G1 no EO (41% vs. 36%) G3 CPE compared with G1 no CPE (36% vs. 41%) No statistically significant effect of G2 EO (OR = 1.13) nor CPE (OR = 0.88) on appropriateness 95% CI: EO: 0.52-2.45; CPE: 0.41-1.91 p=NR	Clustering of visits and baseline appropriateness GEE model

Table E-5. Key question 2 studies first outcome (continued)

Author, Year	Groups	Outcome #1, Exact Measure Used	Timing of Measurement, Data Source	N analyzed for This Outcome	Results by Group	Differences in Groups	Covariates Controlled for in Analysis, Statistical Methods Used
Wetter et al., 2006[37]	G1: Single standard telephone-counseling session (increase reach) G2: Multiple enhanced telephone counseling sessions (multi-component)	Health-related decisions or behavior (applicable for general public/patients) Smoking abstinence: self-report of no smoking during the previous 7 days	5- and 12-week follow-up assessments Self-report	NR	% abstinent Week 5: G1: 11.7% G2: 20.3% Week 12: G1: 20.5% G2: 27.4%	Treatment effect was significant Difference: OR = 3.8 95% CI: NR p=0.048 G1 vs. G2 Week 5: 8.6% Week 12: 6.9%	Time; demographic and tobacco-related variables Generalized linear mixed model regression
Wolters et al., 2005[38]	G1: Control mailed guidelines (increase reach) G2: Intervention involving package for learning, supporting materials, decision tree, and information leaflets for patients (multi-component)	Behavior (applicable for clinicians) Adherence to guidelines. Appropriate request of PSA. Classified patients in terms of those that met certain indications. Number of PSA ordered in patients with and without indications	Up to 1 year post-intervention Prospective recording of patient data and management immediately after consultation with eligible patient	Patient With Indications N= 69 G1: 39 G2: 30 Patients Without Indications N = 118 (n not reported by groups)	Patient With Indications who had PSA's ordered (in line with guideline) G1: 22, 66.7% G2: 15, 50% Patient w/o indications who had PSA's ordered (non-adherence with guideline) G1: 53.6% G2: 37.1%	Patients w/ no indications G1 vs. G2: 16.7% Chi sq p=00.16) People w/o indications G1 vs. G2: 16.5% Chi-sq p=00.07	Age, group allocation, IPSS and BS Chi square

Table E-5. Key question 2 studies first outcome (continued)

Author, Year	Groups	Outcome #1, Exact Measure Used	Timing of Measurement, Data Source	N analyzed for This Outcome	Results by Group	Differences in Groups	Covariates Controlled for in Analysis, Statistical Methods Used
Wright et al., 2008[39]	G1: Standardized lecture by expert opinion leader (increase motivation) G2: Standardized lecture by expert opinion leader + academic detailing and a toolkit (multi-component)	Clinical outcomes (applicable for general public/patients) Mean # of lymph nodes assessed in patients with stage II colon cancer	360 days before intervention, 360 days after intervention NR	NR	G1: Mean # of nodes assessed: 14.9 G2: Mean # of nodes assessed: 18.1	Difference between G1 and G2 in mean # of nodes: 3.2 Difference: Significant increase in the mean # of lymph nodes assessed and the proportion of cases with 12 or more lymph nodes retrieved for G1 and G2 95% CI: NR p=0.001 No additional increase was found when the opinion leader received academic detailing and the toolkit (G2)	NR Logistic regression

Abbreviations: AF = audit and feedback; AHR = airway hyper-responsiveness; ANCOVA = Analysis of covariance; ANOVA = ANalysis Of Variance; b/t = between; BP = blood pressure; BS = Bother score; CAL = computer-assisted learning; CaP = Cancer of the Prostate; CHF = congestive heart failure; CI = confidence interval; CME = continuing medical education; COPE = Compassionate Options for Progressive Eldercare; CPE = continuing professional education; CRL = control; Ctrl = control; d.f. = degrees of freedom; DASH = Dietary Approaches to Stop Hypertension; DBP = diastolic blood pressure; DVD = optical disc storage format; ECG = electrocardiogram; EMR = electronic medical record; EN = enteral nutrition; EO = Education Outreach; FEV% = Forced Percentual Expiratory Volume; FOBT = fecal occult blood test; G = group; GEE = generalized estimating equations method; GHP = ;GP = general practitioner; HDL = high-density lipoprotein; HLM = Hierarchical Linear Modeling version 5.04, Scientific Software International; ICC = intracluster correlation coefficient; ICU = intensive care unit; IPSS = International Prostate Symptom Score; IQR = interquartile ratio; IRR = _ incidence rate ratio; ITT = intention to treat; kcal/kg-1 = kilocalorie/kilogram;LDL = low-density lipoprotein; LHA = lay health advisor; LUTS = lower urinary tract symptoms; M = Mean; mcg = micrograms; MH = _; mm Hg = millimeter of mercury; MOD = more of moderate or more vigorous; NR = not reported; OR = odds ratio; OTC = Over the counter; PAR = Stanford 7-Day Physical Activity Recall; PD20 = Bronchial responsiveness; PDA = personal digital assistant; PF = _; PSA = prostate-specific antigen; RE = _; RN = registered nurse; RP = _; SBP = systolic blood pressure; SD = standard deviation; SE = standard error; SEPAR = Spanish Society of Pulmonology; SF-36 = Short Form (36) Health Survey; T/F = true/false; TPV = tailored and targeted print and video; VA = Veterans Administration; vs. = versus; WCB = Workers Compensation Board.

Table E-6. Key question 2 studies with a second outcome

Author, Year	Groups	Outcome #2, Exact measure used	Timing of measurement, Data Source	N analyzed for this outcome	Results by group	Differences in Groups	Covariates controlled for in analysis, Statistical methods used
Banait et al., 2003[2]	G1: Mailed guidelines (increase reach) G2: Educational outreach (Multi-component)	Behavior (applicable for clinicians) Findings at open access endoscopy. Number of endoscopies performed.	7 mths before and after intervention Chart	G1: 56 G2: 57 (ITT analysis)	Pre-intervention Major findings G1: 37.4% G2: 31.1% Minor findings: G1: 24.8% G2: 29.4% Normal G1: 37.8% G2: 39.5% Post-intervention Major Findings: G1:35.5% G2: 31.7% Minor findings G1: 25.4% G2: 24.9% Normal findings: G1: 39.1% G2: 43.4%	No change in the relative proportions of major, minor, and normal findings pre-and post-for either group of practices.	NR Non-parametric

Table E-6. Key question 2 studies with a second outcome (continued)

Author, Year	Groups	Outcome #2, Exact measure used	Timing of measurement, Data Source	N analyzed for this outcome	Results by group	Differences in Groups	Covariates controlled for in analysis, Statistical methods used
Becker et al., 2008[4]	G1: Mailed guideline (Increase clinician reach) G2: Guideline implementation (multi-component, clinicians only) G3: Guideline implementation and motivational counseling directed at patient (multi-component, clinicians and patients)	Health-related decisions or behavior (applicable for general public/patients) Overall physical activity. Measured using the FQPA. The FQPA has satisfactory psychometrical properties and allows a calculation of weighted MET hours per week.	baseline, 6 mth, 12 mth self-report	Patient N baseline = 1378 G1: 479 G2: 489 G3: 410 N 6 mths = 1261 G1: 450 G2: 435 G3: 376 N 12 mths = 1211 G1: 425 G2: 421 G3: 365	6 months G1: M =33.51 G2: M = 36.47 G3: M = 36.29 12 months G1: M = 42.88 G2: M = 46.43 G3: M = 45.93	6 months (author provided odds ratios for groups compared with control only) Mean diff (95% CI) G1 vs. G2: 2.96 (-1.63/7.55) G1 vs. G3: 2.78 (-1.78/7.35) G2 vs. G3: 0.177* p=NR 12 months Mean difference (95% CI) G1 vs. G2: 3.55 (-1.45/8.54) G1 vs. G3: 2.52 (-2.48/7.50) G2 vs. G3: 1.036* p=NR	Sex, age, fear avoidance, physical activity, and number of days in pain during previous 6 months Multilevel mixed modeling accounting for clustering of data on practice level

Table E-6. Key question 2 studies with a second outcome (continued)

Author, Year	Groups	Outcome #2, Exact measure used	Timing of measurement, Data Source	N analyzed for this outcome	Results by group	Differences in Groups	Covariates controlled for in analysis, Statistical methods used
Bekkering et al., 2005[5]	G1: Received guidelines by mail (increase reach) G2: Received guidelines + active training strategy (multi-component)	Clinical outcomes (applicable for general public/patients) Physical functioning (QBPDS), which consists of 20 activities of daily living. Each activity is scored on a 6-point scale ranging from 0 ("no trouble") to 5 ("unable to"), and the total score ranges from 0 ("no dysfunction") to 100 ("maximum dysfunction").	Baseline, 6, 12, 26, and 52 weeks after baseline Self-report	Baseline Overall: 511 patients G1: 259 patients G2: 256 patients 6 weeks Overall: 511 patients G1: 259 patients G2: 256 patients 12 weeks Overall: 511 patients G1: 259 patients G2: 256 patients 26 weeks Overall: 511 patients G1: 259 patients G2: 256 patients 52 weeks Overall: 511 patients G1: 259 patients G2: 256 patients	Mean scores and interquartile ranges Baseline G1: 40.5 (26.3-55.8) G2: 38.0 (26.5-50.5) 6 weeks G1: 23.5 (11.0-37.8) G2: 24.0 (13.0-40.0) 12 weeks G1: 17.5 (6.0-30.8) G2: 20.0 (7.0-32.8) 26 weeks G1: 11.0 (4.0-29.0) G2: 16.0 (5.0-32.0) 52 weeks G1: 13.0 (4.8-29.0) G2: 17.0 (4.6-32.0)	Adjusted absolute differences (G2-G1): 6 weeks: 1.96 (-1.44 to 5.37) 12 weeks: 2.83 (-0.66 to 6.31) 26 weeks: 4.00 (0.68 to 7.33) 52 weeks: 3.55 (-0.25 to 7.35)	Sex, Previous episode of back pain, duration of current episode of back pain, pain and coping inventory relaxation subscale. Clustering of practices, physical therapists, patients, time points. Multilevel modeling; Wald chi-square tests

Table E-6. Key question 2 studies with a second outcome (continued)

Author, Year	Groups	Outcome #2, Exact measure used	Timing of measurement, Data Source	N analyzed for this outcome	Results by group	Differences in Groups	Covariates controlled for in analysis, Statistical methods used
Bishop and Wing, 2006[6]	G1: Control (not abstracted) G2: Physician only (increase reach) G3: Physician and patient (multi-component)	Behavior (applicable for clinicians) Guideline-concordant treatment advice for 5-12 week post injury period. The compulsory WCB physician report forms were collected and scored. Dichotomous measure of 1 = presence of concordant/ Discordant behavior.	Once at 5-12 weeks Workers' Compensation Board reports	5-12 weeks Overall N: 448 G2: 154 G3: 145	Concordant Supervised exercise program: G2: 19% G3: 18% Return to work G2: 24% G3: 23% Ref to Interdisc. Program: G2: 4% G3: 0% Discordant behavior Physiotherapy G2: 41% G3: 42% NOTE: Authors did not provide any figures, tables, or data for the >12 week measures. Only state no change seen in the recommended use of ongoing supervised exercise programs.	Only compared control to each condition Supervised exercise G2 vs. G3: 1%*, p=NR Return to work G2 vs. G3: 1%*,p=NR Referred to Interdisc program G2 vs. G3: 4%*,p=NR Physiotherapy G2 vs. G3: 1%*, p=NR NOTE: Authors did not analyze between groups difference from each other. Only state no change seen in the recommended use of ongoing supervised exercise programs.	None Chi-square

Table E-6. Key question 2 studies with a second outcome (continued)

Author, Year	Groups	Outcome #2, Exact measure used	Timing of measurement, Data Source	N analyzed for this outcome	Results by group	Differences in Groups	Covariates controlled for in analysis, Statistical methods used
Campbell et al., 2004[7]	G1: Control (not abstracted) G2: LHA (increase motivation) G3: TPV (multi-component) G4: TPV and LHA (multi-component)	Health-related decisions or behavior (applicable for general public/patients) Physical activity. 16-item checklist assessed frequency of different types of activity, with response options of "don't do", "1-3 times/month" "1-2/week" "3-4/week" "5 or more/week" Total physical activity was the sum of all 16 items. Moderate-vigorous recreational activity was calculated by summing responses for 11 of the items (walking, jogging, swimming, biking, sports, etc.). PA was then calculated in terms of MET hours per week. MET hours/week calculated by multiplying frequency time duration (converted	Baseline and 1 yr follow-up Self-report	N = 587 G2: 123 G3: 159 G4: 176	Recreational activity MET Baseline G2: 10.5 (0.90) G3: 9.5 (0.80) G4: 9.7 (0.76) Followup G2: 10.6 (0.70) G3: 10.9 (0.61) G4: 9.7 (0.60) % meeting PA recommendations Baseline G2: 45.5 G3: 41.1 G4: 40.9 Followup G2: 43.9 G3: 46.3 G4: 45.9	Recreational activity MET Baseline G2 vs. .G3: 1.0* G2 vs. G4: 0.8* G3 vs. G4: 0.2* ns, p=0.80 Followup G2 vs. G3: 0.3* G2 vs. G4:0.9* G3 vs. G4: 1.2* p=0 .04 for TPV intervention versus control Baseline G2vsG3: 4.4* G2vsG4: 4.6* G3vsG4: 0.2* ns, p=0.68 Followup G2vsG3: 2.4* G2vsG4: 2.0* G3vsG4: 0.4* p=0.04 for the TPV "intervention main effect" (see outcome 1)	Demographics regression models

Table E-6. Key question 2 studies with a second outcome (continued)

Author, Year	Groups	Outcome #2, Exact measure used	Timing of measurement, Data Source	N analyzed for this outcome	Results by group	Differences in Groups	Covariates controlled for in analysis, Statistical methods used
Campbell et al., 2004[7] (continued)		to hours per week) by the MET value for each activity and summing across items.					
Carney et al., 2005[8]	G1: Mailed health information (increase reach) G2: Telephone counseling (increase motivation)	Health-related decisions or behavior (applicable for general public/patients) Mean time interval between screening exams (measured in months).	Designated time interval for up-tp-date status was within 14 months of the intervention date Objective measurement; NIH mammography registry	Overall N: 258 G1: 126 G2: 132	M (SD) between 1st and 2nd intervention G1: 30.3 (15.9) G2: 25.7 (14.4) M (SD) 15 months after 2nd intervention G1: 33.2(19.4) G2: 25.8 (16.4)	Difference in groups between 1st and 2nd intervention, 4.6*, P =0.02 Difference in groups 15mth after 2nd intervention: 7.4*, p=0 .001	NR t-test

Table E-6. Key question 2 studies with a second outcome (continued)

Author, Year	Groups	Outcome #2, Exact measure used	Timing of measurement, Data Source	N analyzed for this outcome	Results by group	Differences in Groups	Covariates controlled for in analysis, Statistical methods used
Christakis et al., 2006[9]	G1: Usual care (not abstracted) G2: Parental content Alone (increase reach) G3: Provider notification alone (not abstracted) G4: Parental content and provider notification (multi-component)	Health-related decisions or behavior (applicable for general public/patients) Patients were asked about their preventative practices. In some cases, there was >1 question for each behavior. For example, for smoking, they asked whether patient had quit, had set a quit date, or had contacted the tobacco quitline, all of which were associated with successful smoking cessation.	2 to 4 weeks after the scheduled well-child visit self-report	Unclear	IRR (95%CI) G2: 1.05(1.01-1.09) G4: 1.07 (1.03-1.11)	G4 and G2 significantly differed from G1 (reference) G2 vs. G4: 0.02	NR Poisson Regression

Table E-6. Key question 2 studies with a second outcome (continued)

Author, Year	Groups	Outcome #2, Exact measure used	Timing of measurement, Data Source	N analyzed for this outcome	Results by group	Differences in Groups	Covariates controlled for in analysis, Statistical methods used
Davis et al., 2004[10]	G1: Control - guidelines by mail (increase reach) G2: Intermediate (multi-component) G3: High intervention (multi-component)	Clinical outcomes (applicable for general public/patients) Epilepsy-specific QOL. Mastery = measures their sense of mastery of their illness Impact = the impact of epilepsy on patients' lives	Baseline and 12 month follow-up Self-report	Patients at Baseline: Overall:1,133 G1: 370 G2: 364 G3: 399 Patients at follow-up Overall: 811 G1: 255 G2: 269 G3: 287	Baseline Mastery G1: 20.1 (19.4/20.8) G2: 20.2 (19.7/20.7) G3: 19.9 (19.2/20.7) Impact G1: 28.4 (27.1/29.6) G2: 29.1 (28.1/30.2) G3: 27.8 (26.0/29.7) 12mth follow-up Mastery G1: 20.3 (19.7/20.8) G2: 20.5 (19.9/21.0) G3: 19.7 (19.1/20.4) Impact G1: 28.8 (27.6/29.9) G2: 29.4 (28.3/30.5) G3: 28.1 (26.3/30.0)	No significant differences in scale scores were seen across the arms at baseline or after the intervention Mastery G1 vs. G2: 0.1* G1 vs. G3: 0.2* G2 vs. G3: 0.3* p=NR Impact G1 vs. G2: 0.7* G1 vs. G3: 0.6* G2 vs. G3: 1.3* p=NR 12 month follow-up Mastery G1 vs. G2: 0.2* G1 vs. G3: 0.6* G2 vs. G3: 0.8* p=NR Impact G1 vs. G2: 0.6* G1 vs. G3: 0.7* G2 vs. G3: 1.3* p=NR	Deprivation, age, sex, and the training status of the practice t tests

Table E-6. Key question 2 studies with a second outcome (continued)

Author, Year	Groups	Outcome #2, Exact measure used	Timing of measurement, Data Source	N analyzed for this outcome	Results by group	Differences in Groups	Covariates controlled for in analysis, Statistical methods used
Elder et al., 2005[12]	G1: Culturally targeted print-materials + activity inserts (increase reach) G2: Tailored print materials + activity inserts + supporting materials (multi-component). G3: Tailored print materials + in-person promotora (multi-component)	Clinical outcomes (applicable for general public/patients) Total dietary fiber (g)	Baseline and 12 week follow-up Self-report face-to-face interview	Follow-up N = 313 G1: 107 G2: 99 G3: 107	Adjusted Mean at Time 2 G1: 15.6g G2: 17.2g G3: 16.1g	G1 vs. G2: 1.6* G1 vs. G3: 0.5* G2 vs. G3: 1.1* p=NR, but not significant	baseline mean Tukey-Kramer multiple comparison test
Feldstein et al., 2006[13]	G1: Usual care (not abstracted) G2: EMR reminder (increase reach for clinicians) G3: EMR reminder and patient reminder (via letter with educational materials (multi-component)	Health-related decisions or behavior (applicable for general public/patients) Percent receiving Bone mineral density measurement via DXA	Within 6 months of intervention Electronic data provided by referral site	G1: 101 G2: 101 G3: 109	G1: .9% G2: 23.8% G3:22.9%	Difference: G2 vs. G1 22.9 95% CI: .39 (.28-.50) p=NR G3 vs. G1 22 95% CI: .31 (.21-.43) p=NR G3 vs. G2 -.9 95% CI:NR	Fracture type, age, weight less than 127 pounds, osteoporosis diagnosis, and Charlson co-morbidity index. General linear modeling using treatment group, fracture type, age, weight, osteoporosis diagnosis and Charlson Comorbidity Index indicators. Models include independent variables

Table E-6. Key question 2 studies with a second outcome (continued)

Author, Year	Groups	Outcome #2, Exact measure used	Timing of measurement, Data Source	N analyzed for this outcome	Results by group	Differences in Groups	Covariates controlled for in analysis, Statistical methods used
Feldstein et al., 2006[13] (continued)							significant in univariate analyses at p<.10. Continuous outcome measures change scores regressed on the baseline values and indicators of treatment groups. Logistic regression used for unadjusted results

Table E-6. Key question 2 studies with a second outcome (continued)

Author, Year	Groups	Outcome #2, Exact measure used	Timing of measurement, Data Source	N analyzed for this outcome	Results by group	Differences in Groups	Covariates controlled for in analysis, Statistical methods used
Gattellari et al., 2005[14]	G1: Leaflet (increase reach) G2: Video (increase reach) G3: Booklet (increase reach)	Behavioral intentions to use or apply the evidence Propensity to undergo screening during the next 12 months (5-point likert ranging from "definitely not want", "unlikely to want", "not mind", "probably want" to "definitely want"	21 days after receiving information (range 15-31) Self-report	N=405	Post-test: n (%) G1: Definitely not want: 3 (2.2) Unlikely to want: 17 (12.5) Not mind: 28 (20.6) Probably want: 46 (33.8) Definitely want: 42 (30.9) G2: Definitely not want: 6 (4.3) Unlikely to want: 17 (12.3) Not mind: 21 (15.2) Probably want: 41 (29.7) Definitely want: 53 (38.4) G3: Definitely not want: 6 (4.6) Unlikely to want: 18 (13.7) Not mind: 26 (19.8) Probably want: 44 (33.6) Definitely want: 37 (28.2)	Absolute difference in propensity to go screening: G2-G1: definitely not want: +2.1%* be unlikely to want: -0.2%* not mind: -5.4%* probably want: -4.1%* definitely want: +7.5%* G3-G1: definitely not want: +2.4%* be unlikely to want: +1.2%* not mind: -0.8%* probably want: -0.2%* definitely want: -2.7%* any difference between g1, g2, g3: p=0.31	None McNemar's chi-squared

Table E-6. Key question 2 studies with a second outcome (continued)

Author, Year	Groups	Outcome #2, Exact measure used	Timing of measurement, Data Source	N analyzed for this outcome	Results by group	Differences in Groups	Covariates controlled for in analysis, Statistical methods used
Hagmolen et al., 2008[15]	G1: Guideline dissemination (increase reach) G2: Guideline dissemination + educational program (increase ability) G3: Guideline dissemination + educational program + individualized treatment advice based on airway responsiveness and symptoms (multi-component)	Clinical outcomes (applicable for general public/patients) Changes in asthma symptom scores Total Symptom Score: Mean score per day. Cough, wheeze, and shortness of breath were scored twice daily (0=no complaints, 1=once a day, 2=more than once a day, 3=whole day) in a two week diary. Range = 0-18. Calculated total symptom score, night symptom score, and the number of symptom-free days	2-week diary ; respondents entered scores 2 times a day (morning and night) for 2 weeks Self-report	G1: n = 98 G2: n = 133 G3: n = 131	Total Symptom Score: G1: M = 0.9 (SE = 0.2) G2: M = 1.2 (SE= 0.2) G3: M = 1.0 (SE = 0.2) Post-hoc analysis: G1&G2: 1.1 (SE = 0.1) G3: 1.0 (SE=0.2) Nocturnal symptom score: G1: M = 0.3 (SE = 0.1) (difference between baseline and end of study = -0.24) G2: M = 0.5 (SE= 0.1) (difference between baseline and end of study = -0.07) G3: M = 0.4 (SE = 0.1) (difference between baseline and end of study = -0.15) Post-hoc analysis: G1&G2: 1.1 (SE = 0.1) G3: 1.0 (SE=0.2)	Total Symptom Score: Difference: G1 vs. G2: 0.3* G1 vs. G3: 0.1* G2 vs. G3: 0.2* No significant differences between all 3 groups p=0 .08 Significant difference between baseline and end of study measurement in Groups 1 (-.6, p < .05) and G3 (-.5, P<0.05) Post-hoc analysis: G1&G2 vs. G3: 0.1* Significant difference between groups p=0 .6 Significant difference between baseline and end of study measurement in Groups 1&2 (-.4, p < .05) and G3 (-.5, P<0.05) Nocturnal symptom score: G1 vs. G2: 0.2* G1 vs. G3: 0.1* G2 vs. G3: 0.1*	NR Mixed model ANOVA analyses

Table E-6. Key question 2 studies with a second outcome (continued)

Author, Year	Groups	Outcome #2, Exact measure used	Timing of measurement, Data Source	N analyzed for this outcome	Results by group	Differences in Groups	Covariates controlled for in analysis, Statistical methods used
Hagmolen et al., 2008[15] (continued)						Significant overall treatment effect for all 3 groups. p=0.02. Post-hoc analysis G1&G2 vs. G3: 0.1* No significant difference p=0.2	
Jain et al., 2006[16]	G1: Passive intervention-guidelines by mail (increase reach) G2: Active intervention (multi-component)	Clinical outcomes (applicable for general public/patients) Glycemic control Measured 3 different ways: Daily average glucose % of ICU stay with glucose between 4.4-6.1 mmol/L Hyperglycemic index above 6.1	Baseline and 12 month follow-up Observation	Practice Overall: 58 ICUs randomized as 50 clusters G1: 25 clusters G2: 25 clusters Patients Baseline Overall: 623 G1: 298 G2: 325 Follow-up Overall: 612 G1: 305 G2: 307 Note: patients not the same at baseline and follow-up. Authors took a cross-sectional survey at both time points.	Daily Average Glucose (raw Median with interquartile ranges) Baseline G1:8.2 (7.2/9.5) G2: 8.1 (7.3/9.7) Follow-up G1: 8.1 (7.1/9.4) G2: 7.7 (6.9/8.8) % of ICU, Median Baseline G1: 5.9 (0.0/19.0) G2: 3.4 (0.0/14.8) Follow-up G1: 7.7 (0.7/22.6) G2: 13.5 (3.6/27.9) Hyperglycemic index, median Baseline G1: 2.1 (1.2/3.5) G2: 2.1 (1.3/3.8) Follow-up G1: 2.0 (1.1/3.4) G2: 1.7 (0.9/2.7)	Difference (G1 minus G2) in change: Daily Average Glucose p = .003 % of ICU p=0 .003 Hyperglycemic index p=0.003	NR Linear mixed model

Table E-6. Key question 2 studies with a second outcome (continued)

Author, Year	Groups	Outcome #2, Exact measure used	Timing of measurement, Data Source	N analyzed for this outcome	Results by group	Differences in Groups	Covariates controlled for in analysis, Statistical methods used
Kennedy et al., 2003[19]	G1: Control (not abstracted) G2: Information (increase reach) G3: Interview (increase motivation)	Knowledge about the evidence "I fully understand what treatment options are available to me: strongly agree, agree, not sure, disagree, strongly disagree	Post-consultation Self-report	Overall: 717 G2: 244 G3: 236	Strongly Agree: G2: 86 (35.7%) G3: 71 (30.7%) Agree G2: 101 (41.9%) G3: 120 (51.9%) Not sure G2: 27 (11.2%) G3: 20 (8.7%) Disagree G2: 23 (9.5%) G3: 17 (7.4%) Strongly disagree: G2: 4 (1.7%) G3: 3 (1.3%)	NR	Consultant sex age baseline knowledge Ordinal regression
King et al., 2007[20]	G1: Attention control (not abstracted) G2: Counselor via phone (increase motivation) G3: Automated counselor via phone (increase reach)	Clinical outcomes (applicable for general public/patients) The CHAMPS physical activity questionnaire for older adults used to supplement the PAR. Estimates of mean times per week engaged in 30 minutes or more of MOD physical activity and mean minutes per week in MOD activity can also be derived from the CHAMPS.	Baseline, 6, 12 months Self report	N = 189 G2: 66 G3: 61	CHAMPS kcal/kg-1/day-1 (SD) Baseline G2: 1.5 (1.8) G3: 1.4 (1.5) 6 months-baseline Δ G2: 2.1 (2.4) G3: 1.3 (2.5) 12 months-baseline Δ G2: 2.1 (2.6) G3: 2.0 (3.0) CHAMPS min. of MOD+ activity/week Baseline G2: 166.1 (210.9) G3: 154.0 (164.0)	Difference in Δ scores at 6 months CHAMPS kcal/kg-1/day-1 (SD) G2 vs. G3: 0.8*, p=NR CHAMPS min. of MOD+ activity/week G2 vs. G3: 78.8*, p=NR CHAMPS days/week engaged in ≥ 30 min of MOD+ G2 vs. G3:0.5*, p=NR Difference in Δ scores at 12 months CHAMPS kcal/kg-1/day-1 (SD) G2 vs. G3=0.1*, p=NR	Baseline levels of dependent variables Gender ANCOVA

Table E-6. Key question 2 studies with a second outcome (continued)

Author, Year	Groups	Outcome #2, Exact measure used	Timing of measurement, Data Source	N analyzed for this outcome	Results by group	Differences in Groups	Covariates controlled for in analysis, Statistical methods used
King et al., 2007[20] (continued)		CHAMPS is expected to find higher numbers than the PAR measures because it involves self-reporting of "usual activity levels" over the previous 4 week period			6 months-baseline Δ G2: 217.3 (252.3) G3: 138.5 (258.0) 12 months-baseline Δ G2: 216.7 (272.2) G3: 205.0 (323.9) CHAMPS days/week engaged in ≥ 30 min of MOD+ Baseline G2: 2.5 (2.8) G3: 3.1 (3.8) 6 months-baseline Δ G2: 1.4 (5.7) G3: 0.9 (5.7) 12 months-baseline Δ G2: 5.3 (6.1) G3: 4.7 (5.9)	CHAMPS min. of MOD+ activity/week G2 vs. G3: 11.7*, p=NR CHAMPS days/wk engaged in ≥ 30 min of MOD+ G2 vs. G3:0.6*, p=NR	
Laprise et al., 2009[21]	G1: CME (increase ability) G2: CME + practice enablers and reinforcers (multi-component)	Behavior (applicable for clinicians) Adherence to specific recommendations. Patients considered undermanaged at baseline if no record, for at least 1 recommendation of a preventive action undertaken by their GP in the 12 months prior to the first visit	Baseline and followup (exact time not specified) Retrospective audit information	G1: 948 G2: 1396	Recommendation of antiplatelets # undermanaged at baseline G1: 367 G2: 494 # of patient with recommendation at follow-up G1: 136 (37.1%) G2: 235 (47.6%)	Antiplatelets OR:1.50 (1.00-2.24) Angiotensine OR: 2.19 (1.45-3.30) Lipid-lowering OR:1.50 (0.99-2.30) Beta-blockers OR:1.12 (0.57-2.18)	NR Logistic regression

Table E-6. Key question 2 studies with a second outcome (continued)

Author, Year	Groups	Outcome #2, Exact measure used	Timing of measurement, Data Source	N analyzed for this outcome	Results by group	Differences in Groups	Covariates controlled for in analysis, Statistical methods used
Laprise et al., 2009[21] (continued)		following recruitment. Recorded as binary outcome (present, not present)			Recommendation of Angiotensine converting enzyme inhibitor # undermanaged at baseline G1: 600 G2: 875 # of patient with rxn at follow-up G1: 66 (11.0%) G2: 179 (20.5%) Recommendation of lipd-lowering agent when LDL >2.5 mmol/L # undermanaged at baseline G1: 224 G2: 345 # of patient with recommendation at follow-up G1: 58 (25.9%) G2: 119 (34.5%) Recommendation of beta-blockers in post-MI patients # undermanaged at baseline G1: 110 G2: 143		

Table E-6. Key question 2 studies with a second outcome (continued)

Author, Year	Groups	Outcome #2, Exact measure used	Timing of measurement, Data Source	N analyzed for this outcome	Results by group	Differences in Groups	Covariates controlled for in analysis, Statistical methods used
Laprise et al., 2009[21] (continued)					# of patients with recommendation at follow-up G1: 17 (15.5%) G2: 24 (16.8%)		
Lien et al., 2007,[22] Svetkey et al., 2003,[23] Young et al., 2009[24]	G1: Advice only (increase reach) G2: Advice + behavioral counseling using established intervention (multi-component) G3: Established intervention + DASH dietary recommendations (multi-component)	Clinical outcomes (applicable for general public/patients) Change in weight measured using a calibrated scale	Measured at baseline and 6 months Objective measurement	N= 713 G1: 242 G2: 238 G3: 233	G1: -1.1 (3.2) G2: -4.9 (5.5) G3: -5.8 (5.8)	G2-G1: -3.8, P<0.001 G3-G1: -4.7, P<0.001 G3-G2: -0.9, P =0.07	None Mantzel-Haenzel chi-squared

Table E-6. Key question 2 studies with a second outcome (continued)

Author, Year	Groups	Outcome #2, Exact measure used	Timing of measurement, Data Source	N analyzed for this outcome	Results by group	Differences in Groups	Covariates controlled for in analysis, Statistical methods used
Marcus et al., 2009[25]	G1: Contact control treatment delayed group (not abstracted) G2: Telephone-based individualized feedback (increase motivation) G3: Print-based individualized feedback (increase reach)	Self-efficacy to use the evidence Exercise-specific self-efficacy measured by questionnaire developed by Marcus, et al.	Baseline, 6 and 12 months	NR	G1: 6 Months: 2.47; 12 Months: 2.37 G2: 6 Months: 3.04; 12 Months: 2.86 G3: 6 Months: 2.87; 12 Months: 2.98	Difference: 6 Months: F=10.33; 12 Months: F=18.00 95% CI: NR 6 Months: P<0.0001; 12 Months: P<0.0001	Yes Analysis of covariance, adjusted for treatment effects for gender and seasonal differences. When overall test of between-groups differences was significant at the >05 level, the source of these differences was examined further using single-degree-of-freedom contrasts that compared the active treatment arms with each other as well as with the treatment delayed group.

Table E-6. Key question 2 studies with a second outcome (continued)

Author, Year	Groups	Outcome #2, Exact measure used	Timing of measurement, Data Source	N analyzed for this outcome	Results by group	Differences in Groups	Covariates controlled for in analysis, Statistical methods used
Murtaugh et al.,2005[27]	G1: Usual care (not abstracted) G2: Basic intervention e-mail reminder (increase reach) G3: Augmented intervention of e-mail reminder + package of supporting materials (multi-component)	Discussions about the evidence % giving patients instruction about fluid weight gain	Chart-review of subsequent RN visit, within 45 days of initial intake Chart	354	Overall N: 354 G1: 20.6% G2: 29.9% G3: 39.7%	Difference G2-G1: 9.3%, p=0.097 Difference G3-G1: 19.1%, p=0.001 Difference G3-G2: 9.8%*, CI and p=NR	Sociodemo-graphic variables of the RN (age, gender, race/ethnicity), Rn employment status, educational level and caseload; average baseline characteristics of patients care for by each RN including health, functional status; geographic area where nurse provided care Predictive multivariate modelling

Table E-6. Key question 2 studies with a second outcome (continued)

Author, Year	Groups	Outcome #2, Exact measure used	Timing of measurement, Data Source	N analyzed for this outcome	Results by group	Differences in Groups	Covariates controlled for in analysis, Statistical methods used
Paradis et al.,2011[28]	G1: Paper handouts (increase reach) G2: Educational DVD (increase reach)	Self-efficacy to use the evidence Infant care self-efficacy; assessed using 20 items from the 52-item Infant Care Survey. These included knowledge items such as recognizing gas pains and knowing regular breathing sounds of babies, and skill items such as treating diaper rash and taking the baby's temperature. Each item was rated on a 5-point scale, from 1 (very little confidence) to 5 (quite a lot of confidence).	2 weeks post-intervention Self-report	Overall N: 137 G1: 64 G2: 70	Mean change in Self-Efficacy (from baseline): Overall self-efficacy: G1: 0.14 (SD = 0.26) G2: 0.16 (SD = 0.32) NOTE: baseline self-efficacy G1: 4.6, G2: 4.6 Very confident, n (%): Bathing your baby: G1: 52 (77.6%) G2: 65 (92.9%) Knowing regular breathing sounds of babies: G1: 40 (59.7%) G2: 50 (71.4%) Recognizing congestion: G1: 35 (52.2%) G2: 49 (70.0%) Relieving gas pains: G1: 38 (56.7%) G2: 43 (61.4%) Soothing your crying baby: G1: 46 (68.7%) G2: 55 (78.6%) Breast- or bottle-feeding your baby: G1: 54 (80.6%) G2: 62 (88.6%)	Overall self-efficacy: G2-G1: +0.02, p=0.60 Bathing your baby: G2-G1: 15.3%, p=0.01 Knowing regular breathing sounds of babies: G2-G1: 11.7%, p=0.15 Recognizing congestion: G2-G1: 17.8%, p=0.03 Relieving gas pains: G2-G1: 4.7%, p=0.58 Soothing your crying baby: G2-G1: 9.9%, p=0.19 Breast- or bottle-feeding your baby: +W4 8%, p=0.20	Hispanic ethnicity, babies born at outside hospital, #exclusively breast fed Multivariate regression analysis

Table E-6. Key question 2 studies with a second outcome (continued)

Author, Year	Groups	Outcome #2, Exact measure used	Timing of measurement, Data Source	N analyzed for this outcome	Results by group	Differences in Groups	Covariates controlled for in analysis, Statistical methods used
Partin et al., 2004[29]	G1: Usual care (not abstracted) G2: Pamphlet (increase reach) G3: Video (increase reach)	Behavioral intentions to use or apply the evidence Screening intention was assessed from a single yes/no question regarding whether the patient thought they would have a PSA test in the next year.	1 week post-target appointment Self-report	N = 893 G2: 295 G3: 308	Unadjusted proportions G2: 0.64 G3: 0.61 Adjusted proportions G2: .65 G3: .63	Unadjusted: G2 vs. G3: 0.03*, p=NR Adjusted G2 vs. G3: 0.02* p=NR	Adjusted analysis accounted for marital status, education, race, health status, comorbid conditions, experience with prostate problems, symptom severity, medication use Logistic regression

Table E-6. Key question 2 studies with a second outcome (continued)

Author, Year	Groups	Outcome #2, Exact measure used	Timing of measurement, Data Source	N analyzed for this outcome	Results by group	Differences in Groups	Covariates controlled for in analysis, Statistical methods used
Rebbeck et al., 2006[31]	G1: Dissemination of guidelines by mail (increase reach) G2: Implementation group (multi-component)	Knowledge about the evidence Measured using a questionnaire developed for this study. Questions included: self-rating of knowledge of the guidelines, treatments currently used to manage whiplash, treatments understood to be evidence-based, when and why physiotherapists refer to other disciplines, treatment goals set for whiplash patients, reporting responsibilities, and understanding of yellow flags (see Appendix 1). Score ranges from 0 to 28, with higher scores indicating greater knowledge of the guidelines.	Baseline and 12 months Self-report	Baseline: Overall: 27 G1: 13 G2: 14 After study (12 mo follow-up) Overall: 26 G1: 12 G2: 14	Total knowledge score: Baseline G1: M = 14.6 (SD=2.3) G2: M = 13.6 (SD=3.2) 12 month follow-up G1: 12.8 (SD=3.3) G2: 17.9 (SD=3.5)	Absolute differences: Baseline: G1 vs. G2: 1.0* 12 month follow-up: G1 vs. G2: 5.1* Difference: Physiotherapists in the implementation group increased their knowledge of the guidelines by 5.5 points more than physiotherapists in the dissemination group 95% CI: 2.5-8.4 p=0.001	NR Linear regression, adjusted for before trial score

Table E-6. Key question 2 studies with a second outcome (continued)

Author, Year	Groups	Outcome #2, Exact measure used	Timing of measurement, Data Source	N analyzed for this outcome	Results by group	Differences in Groups	Covariates controlled for in analysis, Statistical methods used
Rimer et al., 2001[32]	G1: No treatment control/usual care (not abstracted) G2: Tailored print (increase reach) G3: Tailored print + telephone counselling (multi-component)	Awareness of the evidence Perceptions of absolute 10-year and lifetime breast cancer risks between self versus other using verbal and numerical anchors. "How likely are you to get breast cancer in: 1. the next 10 years and 2. your life-time? With 5-pt Likert scale, converted to a percentile. Measured as "over-estimate", accurate in estimates, and under-estimate	12-15 months after baseline interview Self-report	Overall N: 1127 G1: 412 G2: 392 G3: 323	Baseline G1: 305*, 74% G2: 274*, 70% G3: 232*, 72% Yearly- overestimate G1: 309*, 75% G2: 282*, 72% G3: 187*, 58% Yearly- Correctly estimate: G1: 103*, 25% G2: 110*, 28% G3: 136*, 42%	Correctly estimate Yearly: p=0.001 G2-G1: 3%, NS G3-G1: 17%, P<0.05 G3-G2: 14%, P<0.05 Any difference in groups P<0.001	None Pearson chi-squared

Table E-6. Key question 2 studies with a second outcome (continued)

Author, Year	Groups	Outcome #2, Exact measure used	Timing of measurement, Data Source	N analyzed for this outcome	Results by group	Differences in Groups	Covariates controlled for in analysis, Statistical methods used
Simon et al., 2005[33]	G1: Mailed educational materials (increase reach) G2: Individual academic detailing (increase ability) G3: Group academic detailing (increase ability)	Clinical outcomes (applicable for general public/patients) Blood pressure control - blood pressure measurements	Baseline, 1 year follow-up, 2 year follow-up objective measurement	NR	NR	Year 1 Difference: G2 more likely to have systolic blood pressure less than 140 mmHg compared to G1, OR: 0.87 (95% CI: 0.55-1.39) p=ns No difference between G3 and G1, OR: 0.98 (95% CI: .65-1.49)	Differences among individual patients Logistic regression
Soler et al., 2010[34]	G1: Control (not abstracted) G2: Training session on the SEPAR guidelines (increase ability) G3: G2 + portable-device for spirometry (multi-component)	Clinical outcomes (applicable for general public/patients) Use of chest X-rays and arterial blood gas studies (secondary outcome)	NR Chart	G1: 1481, G2: 2119, G3: 5556	Blood gases (phase 2) G1: 41.7% G2: 43.1% G3: 31.6% Chest X-rays (phase 2): G1: 74.6% G2: 74.8% G3: 71%	Absolute Difference in blood gas use: G2-G1: +1.4%, P<0.001 G3-G1: -10.1%, P<0.001 Absolute difference in xrays: G2-G1: +0.2%, P<0.001 G3-G1: -3.6%, P<0.001	Baseline values Logistic regression

Table E-6. Key question 2 studies with a second outcome (continued)

Author, Year	Groups	Outcome #2, Exact measure used	Timing of measurement, Data Source	N analyzed for this outcome	Results by group	Differences in Groups	Covariates controlled for in analysis, Statistical methods used
Sullivan et al., 2010[35]	G1: VA guidelines (increase reach) G2: COPE: web-based education program (increase ability)	Behavior (applicable for clinicians) Frequency of using 4-core management strategies over the earlier 2 months How often did you (0-100%): 1) Agree to prescribe opioids when patients request this? 2) Obtain urine toxicology prior to prescribing? 3) Have patient sign a pain contract (specifying prohibited behavior)? 4) Negotiate a patient treatment agreement (specifying functional goals)?	Baseline and 45-60 days post training Self-report	NR	1) Agree to prescribe opioids when patients request this? G1: Pretest: 43.6% Posttest: 38.0% G2: Pretest: 45.6% Posttest: 37.8% 2) Obtain urine toxicology prior to prescribing? G1: Pretest: 41.8% Posttest: 41.6% G2: Pretest: 39.4% Posttest: 39.9% 3) Have patient sign a pain contract (specifying prohibited behavior)? G1: Pretest: 38.8% Posttest: 41.9% G2: Pretest: 37.9% Posttest: 41.7% 4) Negotiate a patient treatment agreement (specifying functional goals)?	No statistically significant differences between groups Q1 (posttest): G1 vs. G2: 0.2* Q2 (posttest): G1 vs. G2: 1.7* Q3 (posttest): G1 vs. G2: 0.2* Q4 (posttest): G1 vs. G2: 0.9*	NR Independent group t tests; intention-to-treat analyses using the GEE

Table E-6. Key question 2 studies with a second outcome (continued)

Author, Year	Groups	Outcome #2, Exact measure used	Timing of measurement, Data Source	N analyzed for this outcome	Results by group	Differences in Groups	Covariates controlled for in analysis, Statistical methods used
Sullivan et al., 2010[35] (continued)					G1: Pretest: 17.1% Posttest: 20.2% G2: Pretest: 15.5% Posttest: 21.1%		
Watson et al., 2002[36]	G1: Guideline materials by postal mail (increase reach) G2: EO session and guidelines (increase ability) G3: CPE session and guidelines (increase ability) G4: Guidelines + EO and CPE (multi-component)	Knowledge about the evidence 5 knowledge items with 7-point Likert scale: antibiotics can predispose a customer to vaginal thrush; elderly customers should not use OTC anti-fungal preparations; if I recommend an OTC anti-fungal preparation, I will reduce the risk of the infection spreading; women who are pregnant should not use anti-fungal preparations and I only recommend OTC anti-fungal preparations if the customer has a previous diagnosis of vaginal thrush	Baseline and post-intervention but timing not specified Self-report	52 pharmacies at baseline (87%) and 50 (83%) at follow-up	Not presented by group	Difference: No significant changes were shown following either intervention in the five knowledge items. Results summarized but not presented by intervention group; just before and after for all pharmacies.	Unclear NR

Table E-6. Key question 2 studies with a second outcome (continued)

Author, Year	Groups	Outcome #2, Exact measure used	Timing of measurement, Data Source	N analyzed for this outcome	Results by group	Differences in Groups	Covariates controlled for in analysis, Statistical methods used
Wolters et al., 2005[38]	G1: Control mailed guidelines (increase reach) G2: Intervention involving package for learning, supporting materials, decision tree, and information leaflets for patients (multi-component)	Behavior (applicable for clinicians) Provision of patient education materials	Up to 1 year post-intervention Prospective recording of patient data and management immediately after consultation with eligible patient	N = 187 G1: 92 G2: 95	G1: 7.6% G2: 51.6%	G1 vs. G2: 44%* OR: 75.5 (no CI reported)	Age, group allocation, IPSS and BS Logistic regression
Wright et al., 2008[39]	G1: Standardized lecture by expert opinion leader (increase motivation) G2: Standardized lecture by expert opinion leader + academic detailing and a toolkit (multi-component)	Clinical outcomes (applicable for general public/patients) Lymph node removal	360 days before intervention, 360 days after intervention NR	NR	# of lymph nodes removed after lecture G1: 306 G2: 320	G1 vs. G2: 14 Difference: No difference between G1 and G2 95% CI: NR p=0.54	# of lymph nodes retrieved 360 days before the standardized lecture Poisson regression

Abbreviations: ANCOVA = Analysis of covariance; ANOVA = ANalysis Of Variance; BS = Bother score; CHAMPS = Community Healthy Activities Model Program for Seniors; CI = confidence interval; CPE = continuing professional education; DASH = Dietary Approaches to Stop Hypertension; DXA = Dual X-ray absorptiometry;EMR = electronic medical record; EO = Education Outreach; FQPA = Freiburg Questionnaire on Physical Activity; G = group; GEE = generalized estimating equations method; ICU = intensive care unit; IPSS = International Prostate Symptom Score; kcal/kg-1 = kilocalorie/kilogram;LDL = low-density lipoprotein; LHA = lay health advisor; M = Mean; MET = metabolic equivalent take; mmol/L = millimoles/liter; MOD = moderate intensity or more vigorous; mths = months; N = number; NR = not reported; OTC = Over the counter; PA = physician's assistant; QBPDS = Quebec Back Pain Disability Scale; QOL = quality of life; RN = registered nurse; SD = standard deviation; SEPAR = Spanish Society of Pulmonology; TPV = tailored and targeted print and video; WCB = Workers Compensation Board; wk = week.

Table E-7. Key question 2 studies with a third outcome

Author, Year	Groups	Outcome #3, Exact Measure Used	Timing of Measurement, Data Source	N analyzed for This Outcome	Results by Group	Differences in Groups	Covariates Controlled for in Analysis, Statistical Methods Used
Becker et al., 2008[4]	G1: Mailed guideline (Increase clinician reach) G2: Guideline implementation (multi-component, clinicians only) G3: Guideline implementation and motivational counseling directed at patient (multi-component, clinicians and patients)	Clinical outcomes (applicable for general public/patients) Quality of life. Measured with the Euro-Qol and Fear Avoidance Beliefs Questionnaire.	Baseline, 6 month, 12 month Self-report	Patient N baseline = 1378 G1: 479 G2: 489 G3: 410 N 6 months = 1261 G1: 450 G2: 435 G3: 376 N 12 months = 1211 G1: 425 G2: 421 G3: 365	6 months G1: M = 66.85 G2: M = 66.59 G3: M = 67.54 12 months G1: M = 67.65 G2: M = 68.46 G3: M = 70.38	6 months (author provided odds ratios for groups compared with control only) Mean diff (95% CI) G1 vs. G2: -0.25(-2.86/2.36) G1 vs. G3: 0.69 (-1.92/3.30) G2 vs. G3: 0.943* p=NR 12 months Mean diff (95% CI) G1 vs. G2: 0.80 (-1.74/3.34) G1 vs. G3: 2.72(0.19/5.26) G2 vs. G3: 1.919* p=NR	Sex, age, fear avoidance, physical activity, and number of days in pain during previous 6 months

Table E-7. Key question 2 studies with a third outcome (continued)

Author, Year	Groups	Outcome #3, Exact Measure Used	Timing of Measurement, Data Source	N analyzed for This Outcome	Results by Group	Differences in Groups	Covariates Controlled for in Analysis, Statistical Methods Used
Bekkering et al., 2005[5]	G1: Received guidelines by mail (increase reach) G2: Received guidelines + active training strategy (multi-component)	Clinical outcomes (applicable for general public/patients) Pain; measured using an 11-point NRS scale ranging from 0 (“no pain”) to 10 (“very severe pain”).	Baseline, 6, 12, 26, and 52 weeks after baseline Self-report	Baseline Overall: 511 patients G1: 259 patients G2: 256 patients 6 weeks Overall: 511 patients G1: 259 patients G2: 256 patients 12 weeks Overall: 511 patients G1: 259 patients G2: 256 patients 26 weeks Overall: 511 patients G1: 259 patients G2: 256 patients 52 weeks Overall: 511 patients G1: 259 patients G2: 256 patients	Mean scores and interquartile ranges Baseline G1: 7.0 (5.0-8.0) G2: 7.0 (5.0-8.0) 6 weeks G1: 3.0 (2.0-5.0) G2: 3.0 (2.0-5.0) 12 weeks G1: 2.0 (1.0-4.0) G2: 2.0 (1.0-4.0) 26 weeks G1: 1.0 (0.0-4.0) G2: 2.0 (1.0-4.0) 52 weeks G1: 1.0 (0.3-3.0) G2: 2.0 (0.0-4.0)	Adjusted absolute differences (G2-G1): 6 weeks: 0.16 (-0.35 to 0.69) 12 weeks: 0.34 (-0.19 to 0.88) 26 weeks: 0.62 (0.06 to 1.18) 52 weeks: 0.55 (-0.02 to 1.11)	Sex, previous episode of back pain, duration of current episode of back pain, pain and coping inventory relaxation subscale. Clustering of practices, physical therapists, patients, time points. Multilevel modeling; Wald chi-square tests

Table E-7. Key question 2 studies with a third outcome (continued)

Author, Year	Groups	Outcome #3, Exact Measure Used	Timing of Measurement, Data Source	N analyzed for This Outcome	Results by Group	Differences in Groups	Covariates Controlled for in Analysis, Statistical Methods Used
Bishop and Wing, 2006[6]	G1: Control (not abstracted) G2: Physician only (increase reach) G3: Physician and patient (multi-component)	Behavior (applicable for clinicians) Guideline-concordant treatment advice for >12-week post injury treatment period. The compulsory WCB physician report forms were collected and scored. Dichotomous measure of 1 = presence of concordant/ Discordant behavior.	Once during 12-16 weeks Workers' Compensation Board reports	>12 weeks Overall N: 428 G2: 149 G3: 139	NR	NR NOTE: Authors did not analyze between groups difference from each other, nor provide any figures, tables, or data for the >12 week measures.	NR

Table E-7. Key question 2 studies with a third outcome (continued)

Author, Year	Groups	Outcome #3, Exact Measure Used	Timing of Measurement, Data Source	N analyzed for This Outcome	Results by Group	Differences in Groups	Covariates Controlled for in Analysis, Statistical Methods Used
Campbell et al., 2004[7]	G1: Control (not abstracted) G2: LHA (increase motivation) G3: TPV (multi-component) G4: TPV and LHA (multi-component)	Health-related decisions or behavior (applicable for general public/patients) CRC screening. Participants were asked whether they had ever had any CRC screening tests and, if so, ow long ago (< 1 yr, 1-2 yrs, 2-5 yrs, or > 5 yrs).	Baseline and 1 yr follow-up Self-report	N = 587 G2: 123 G3: 159 G4: 176	FOBT test in the past year (%) Baseline G2: 23.5 G3: 19.7 G4: 19.5 Followup G2: 33.4 G3: 36.8 G4: 31.0 Other CRC test in the past year (%) Baseline G2: 19.6 G3: 23.7 G4: 26.4 Followup G2: 25.5 G3: 21.1 G4: 14.9	FOBT test Baseline G2 vs. G3: 3.8 G2 vs. G4: 4.0 G3 vs. G4: 0.2 ns, p=0.36 Followup G2vs.G3: 3.4 G2vs.G4: 2.4 G3 vs.G4: 5.8 ns, p=0.08 Other CRC Baseline G2 vs. G3: 4.1 G2 vs. G4: 6.8 G3 vs. G4: 2.7 ns, p=0.75 Followup G2 vs. G3: 4.4 G2 vs. G4: 10.6 G3 vs. G4: 6.2 p=0.04 but looks like this is in comparison to controls	Demographics regression models

Table E-7. Key question 2 studies with a third outcome (continued)

Author, Year	Groups	Outcome #3, Exact Measure Used	Timing of Measurement, Data Source	N analyzed for This Outcome	Results by Group	Differences in Groups	Covariates Controlled for in Analysis, Statistical Methods Used
Feldstein et al., 2006[13]	G1: Usual care (not abstracted) G2: EMR reminder (increase reach for clinicians) G3: EMR reminder and patient reminder (via letter with educational materials (multi-component)	Behavioral intentions to use or apply the evidence Total caloric expenditure from all activity from the Community Health Activities Model Program for Seniors questionnaire (a self-report physical activity questionnaire for older men and women)	Baseline and 6 months post intervention Patient self-report	G1: 32; G2: 38; G3: 38	G1: Pre: 2,325.7; Post: 1980.9 G2: Pre: 3,082.9; Post: 2312.7; G3: Pre: 2,614.4; Post: 2525.9	Difference: G1 vs. G2: -331.8 G1 vs. G3: -545 G2 vs. G3: -213.2 95% CI: NR p=0.32 treatment and UC	Pre-survey response See Outcome #1

Table E-7. Key question 2 studies with a third outcome (continued)

Author, Year	Groups	Outcome #3, Exact Measure Used	Timing of Measurement, Data Source	N analyzed for This Outcome	Results by Group	Differences in Groups	Covariates Controlled for in Analysis, Statistical Methods Used
Hagmolen et al., 2008[15]	G1: Guideline dissemination (increase reach) G2: Guideline dissemination + educational program (increase ability) G3: Guideline dissemination + educational program + individualized treatment advice based on airway responsiveness and symptoms (multi-component)	Clinical outcomes (applicable for general public/patients) Usage of asthma medication. Number of PPD. For ICS this is mean PPD prescribed in 1 year. For β2 agonist it is mean number of PPD used during the diary period.	NR Objective measurement	Overall N: 362 G1: 98 G2: 133 G3: 131 Also conducted post-hoc analysis where Groups 1 and 2 were combined	ICS G1: M =0.4 (SE=0.05) G2: M = 0.5(SE=0.05) G3: M = 0.6 (SE=0.05) β2 agonist G1: M =0.45 (SE=0.01) G2: M = 0.43(SE=0.08) G3: M = 0.29 (SE=0.08)	ICS G1 vs. G2: 0.1* G1 vs.. G3: 0.2* G2 vs. G3: 0.1* Significant overall treatment effect among all 3 groups. p=0.03 Significant difference between baseline and end of study for G3 (.1, P<0.05) Post-hoc analysis (aggregated groups 1 & 2): G1&G2 vs. G3: .2 Significant difference between groups p=0.02 β2 agonist G1 vs. G2: .02* G1 vs. G3: .16* G2 vs. G3: .14* No significant treatment effect between 3 groups. p=0.2 Significant different between baseline and send of study for G3 (-.24, P<0.05)	NR Mixed model ANOVA analyses

Table E-7. Key question 2 studies with a third outcome (continued)

Author, Year	Groups	Outcome #3, Exact Measure Used	Timing of Measurement, Data Source	N analyzed for This Outcome	Results by Group	Differences in Groups	Covariates Controlled for in Analysis, Statistical Methods Used
Hagmolen et al., 2008[15] (continued)						Post-analysis: G1&G2 vs. G3: .15* No significant difference between groups. p=0.2	
Jain et al., 2006[16]	G1: Passive intervention-guidelines by mail (increase reach) G2: Active intervention (multi-component)	Clinical outcomes (applicable for general public/patients) 3 measures assess Clinical Outcomes. Duration of ICU (reported in days, using the median, and the IQR is reported) Hospital length of stay (reported in median days, IQR. Follow-up was censored at 60 days so true upper quartile is undefined) 28-day mortality rate (reported as n, %)	Baseline and 12 month followup Observation	Practice Overall: 58 ICUs randomized as 50 clusters G1: 25 clusters G2: 25 clusters Patients Baseline Overall: 623 G1: 298 G2: 325 Follow-up Overall: 612 G1: 305 G2: 307 Note: the patients were not the same at baseline and follow-up. The authors took a cross-sectional survey at both time points.	ICU LOS (median, IQR) Baseline G1: 14.9 (8.3/29.9) G2: 14.4 (7.3/32.3) Follow-up G1: 13.7 (7.8/28.5) G2: 13.9 (8.6/33.4) Hospital LOS (median, IQR) Baseline G1: 27.4 (15.3/60) G2: 28.2 (14.4/60) Follow-up G1: 28.8 (15.0/60) G2: 29.1 (14.7/60) 28 day mortality (n, %) Baseline G1: 63 (21.1%) G2: 68 (20.9%) Follow-up G1: 56 (18.4%) G2: 56 (18.2%)	No significant differences in change ICU LOS ΔG1 : -1.2* ΔG2: -0.5* ΔG1-ΔG2: 0.7* p=NR Hospital LOS ΔG1 : 1.4* ΔG2: 0.9* ΔG1-ΔG2: 0.5* p=NR 28 day mortality ΔG1 : -2.7%* ΔG2: -2.7* ΔG1-ΔG2: 0* p=NR	NR Fisher's randomization test of the log-rank statistic

Table E-7. Key question 2 studies with a third outcome (continued)

Author, Year	Groups	Outcome #3, Exact Measure Used	Timing of Measurement, Data Source	N analyzed for This Outcome	Results by Group	Differences in Groups	Covariates Controlled for in Analysis, Statistical Methods Used
Kennedy et al., 2003[19]	G1: Control (not abstracted) G2: Information (increase reach) G3: Interview (increase motivation)	Behavioral intentions to use or apply the evidence Preference at baseline a binary variable was used: 0 = no preference held post-consultation and 1 = preference formed post consultation. For those who did hold a preference at baseline a nominal variable was produced with three categories: preference maintained, preference changed, no preference (the woman no longer held a preference post-consultation).	Baseline and post-consultation. Self-report	Overall: 685 G2: 234 G3: 226	Women with no preference at baseline #,% G2: 135, 57.7% G3: 114, 50.4%	Only reported differences between each group and the control	Consultant sex; Consultant year of qualification; Age; Baseline menorrhagia severity; Baseline knowledge; Previous treatment – D&C; Previous treatment – OCP; Previous treatment – hormonal drugs; Previous treatment – non-hormonal drugs; Duration of problem; Any previous surgery; Baseline preferences (where preference held at baseline); Recruitment period Multinomial logistic regression

Table E-7. Key question 2 studies with a third outcome (continued)

Author, Year	Groups	Outcome #3, Exact Measure Used	Timing of Measurement, Data Source	N analyzed for This Outcome	Results by Group	Differences in Groups	Covariates Controlled for in Analysis, Statistical Methods Used
Lien et al., 2007,[22] Svetkey et al., 2003,[23] Young et al., 2009[24]	G1: Advice only (increase reach) G2: Advice + behavioral counseling using established intervention (multi-component) G3: Established intervention + DASH dietary recommendations (multi-component)	Clinical outcomes (applicable for general public/patients) Percent that met at least 3 health goals	6-months and 18 months Objective measurement	N at 6 months G1: 248 G2: 239 G3: 242 N at 18 months G1: 254 G2: 244 G3: 254	6 months: G1: 11.7% G2: 19.3% G3: 44.6% 18 months G1: 11.0% G2: 11.9% G3: 33.5%	6 months: G2-G1: 7.6%* P<0 .02 G3-G1: 32.9%* P<0 .0001 G3-G2: 25.3%* P<0 .0001 18 months: G2-G1: 0.9%* p=NR, but non-significant G3-G1: 22.5%* P<0 .0001 G3-G2: 21.6%* P< 0.0001	None Mantzel-Haenzel chi-squared

Table E-7. Key question 2 studies with a third outcome (continued)

Author, Year	Groups	Outcome #3, Exact Measure Used	Timing of Measurement, Data Source	N analyzed for This Outcome	Results by Group	Differences in Groups	Covariates Controlled for in Analysis, Statistical Methods Used
Marcus et al., 2009[25]	G1: Contact control treatment delayed group (not abstracted) G2: Telephone-based individualized feedback (increase motivation) G3: Print-based individualized feedback (increase reach)	Behavioral intentions to use or apply the evidence Decision making for exercise measured by decisional balance instrument by Marcus, et al.	Baseline, 6 and 12 months Self-report	NR	G1: 6 Months: -2.95; 12 Months: -3.64 G2: 6 Months: 13.38; 12 Months: -0.75 G3: 6 Months: 15.45; 12 Months: 14.12	Difference: 6 Months: F=4.49; 12 Months: F=6.04 95% CI: NR 6 Months: P<0.0122 12 Months: P<0.0028	Yes Analysis of covariance, adjusted for treatment effects for gender and seasonal differences. When overall test of between-groups differences was significant at the >05 level, the source of these differences was examined further using single-degree-of-freedom contrasts that compared the active treatment arms with each other as well as with the treatment delayed group.

Table E-7. Key question 2 studies with a third outcome (continued)

Author, Year	Groups	Outcome #3, Exact Measure Used	Timing of Measurement, Data Source	N analyzed for This Outcome	Results by Group	Differences in Groups	Covariates Controlled for in Analysis, Statistical Methods Used
Murtaugh et al.,2005[27]	G1: Usual care (not abstracted) G2: Basic intervention e-mail reminder (increase reach) G3: Augmented intervention of e-mail reminder + package of supporting materials (multi-component)	Discussions about the evidence % giving patients instructions about Shortness of breath	chart-review of subsequent RN visit, within 45 days of initial intake Chart	354	Overall N: 354 G1: 18.1% G2: 31.1% G3: 28.9%	Difference G2-G1: 13.0%, p=0.021 Difference G3-G1: 10.8%, p=0.053 Difference G3-G2: -2.2%*, CI and p=NR	Socio-demographic variables of the RN (age, gender, race/ethnicity), Rn employment status, educational level and caseload; average baseline characteristics of patients care for by each RN including health, functional status; geographic area where nurse provided care Predictive multivariate modeling

Table E-7. Key question 2 studies with a third outcome (continued)

Author, Year	Groups	Outcome #3, Exact Measure Used	Timing of Measurement, Data Source	N analyzed for This Outcome	Results by Group	Differences in Groups	Covariates Controlled for in Analysis, Statistical Methods Used
Paradis et al.,2011[28]	G1: Paper handouts (increase reach) G2: Educational DVD (increase reach)	Clinical outcomes (applicable for general public/patients) Health care utilization including the # of additional clinic visits, # of parent-initiated phone calls, total # of professional consultations, proportion with >1 additional visit, and # of emergency dept visits between the enrollment visit and the 2 month well-child visit. “Professional consultations” -- the combination of clinic visits and parent-initiated phone calls. Additional office visits -- any (problem-related) visit outside of the usual well-child schedule	2 months post-enrollment Chart	Overall N: 137 G1: 67 G2: 70	Number of additional clinic visits: G1: 2.0 (SD=1.1) G2: 1.6 (SD=1.2) Number of parent-initiated phone calls: G1: 1.8 (SD=1.9) G2: 1.1 (SD=1.8) Total professional consultations: G1: 4.0 (SD=3.0) G2: 2.9 (SD=2.8) Proportion with >1 additional visit: G1: 42 (63%) G2: 27 (39%) Number of emergency department visits: G1: 0.2 (SD = 0.6) G2: 0.2 (SD = 0.5)	Number of additional clinic visits: G2-G1: - 0.4, 95% CI: -0.80 to -0.01 p=0.05 Number of parent-initiated phone calls: G2-G1: -0.7, 95% CI: -1.22 to -0.01 p=0.05 Total professional consultations: G2-G1: -1.1, 95% CI: -2.00 to -0.03 p=0.04 Proportion with >1 additional visit: G2-G1: -15, 95% CI NR p=0.01 Number of emergency department visits: G2-G1: 0, 95% CI: -0.20 to 0.18 p=0.91	Hispanic ethnicity, babies born at outside hospital, #exclusively breast fed Multivariate regression analysis

Table E-7. Key question 2 studies with a third outcome (continued)

Author, Year	Groups	Outcome #3, Exact Measure Used	Timing of Measurement, Data Source	N analyzed for This Outcome	Results by Group	Differences in Groups	Covariates Controlled for in Analysis, Statistical Methods Used
Partin et al., 2004[29]	G1: Usual care (not abstracted) G2: Pamphlet (increase reach) G3: Video (increase reach)	Discussions about the evidence Patient participation in CaP screening decision making was assessed by a single question about whether CaP screening was discussed at their last clinic visit.	1 week post-target appointment self-report	N = 893 G2: 295 G3: 308	Unadjusted proportions G2: 0.41 G3: 0.35 Adjusted proportions G2: .35 G3: .41	Unadjusted G2 vs. G3: 0.06*, P value NR Adjusted G2 vs. G3: 0.06*, P value NR	Adjusted analysis accounted for marital status, education, race, health status, comorbid conditions, experience with prostate problems, symptom severity, medication use logistic regression

Table E-7. Key question 2 studies with a third outcome (continued)

Author, Year	Groups	Outcome #3, Exact Measure Used	Timing of Measurement, Data Source	N analyzed for This Outcome	Results by Group	Differences in Groups	Covariates Controlled for in Analysis, Statistical Methods Used
Rebbeck et al., 2006[31]	G1: Dissemination of guidelines by mail (increase reach) G2: Implementation group (multi-component)	Clinical outcomes (applicable for general public/patients) Disability due to acute whiplash - measured more specifically using a 5-item adapted version of the Core Outcome Measure for neck pain. Each item was scored on a 5-point response scale. Summation of the 5 items yields a score ranging from 5 to 25; higher scores indicate greater perceived disability.	Baseline, month 1.5, month 3, month 6, month 12 Self-report	Baseline: G1: 28 G2: 71 Month 1.5 G1: 24 G2: 64 Month 3 G1: 23 G2: 59 Month 6 G1: 19 G2: 56 Month 12 G1: 26 G2: 67	Baseline: G1: M=15.5, SD=3.2 G2: 16.3, SD=3.7 Month 1.5 G1: 10.9, SD=3.8 G2: 13.5, SD=4.9 Month 3 G1: 10.2, SD=3.7 G2: 11.5, SD=4.4 Month 6 G1: 9.4, SD=4.3 G2: 10.9, SD=5.2 Month 12 G1: 10.0, SD=4.2 G2: 10.3, SD=4.4	Baseline Difference (G1 vs. G2): 0.3* 95% CI: -2.1 to 2.7 p=0.08 Month 1.5 Difference (G1 vs. G2): 2.8* 95% CI: -0.5 to 6.2 p=0.09 Month 3 Difference (G1 vs. G2): 1.3* 95% CI: -1.3 to 3.8 p=0.31 Month 6 Difference (G1 vs. G2): 1.7* 95% CI: -2.3 to 5.7 p=0.38 Month 12 Difference (G1 vs. G2): 0.3* 95% CI: -2.4 to 3.0 p=0.85	NR t-test, adjusted using methods for cluster-randomized trials

Table E-7. Key question 2 studies with a third outcome (continued)

Author, Year	Groups	Outcome #3, Exact Measure Used	Timing of Measurement, Data Source	N analyzed for This Outcome	Results by Group	Differences in Groups	Covariates Controlled for in Analysis, Statistical Methods Used
Rimer et al., 2001[32]	G1: No treatment control/usual care (not abstracted) G2: Tailored print (increase reach) G3: Tailored print + telephone counselling (multi-component)	Knowledge about the evidence 2 true/false questions on breast CA and mammography (% correct): 1) Mammograms are more effective for women 50-65. 2) Women over 50 are at higher risk for breast cancer	1 week following receipt of intervention Self-report	1127	Mammograms are more effective for women 50-65, % correct: G1: 11% G2: 15% G3: 20% Women over 50 are at higher risk for breast cancer, % correct G1: 25% G2: 27% G3: 37%	Absolute difference in knowledge, mammogram effectiveness: G2-G1: +4%*, p=NS G3-G1: +9%*, P<0.05 G3-G2: +5%*, p=NS Any difference in groups: p=0.007 Absolute difference in knowledge, risk for cancer: G2-G1: +2%*, p NS G3-G1: +12%*, P<0.05 G3-G2: +10%*, P<0.05 Any difference in groups: p=0.001	None Pearson chi-square and F-test
Simon et al., 2005[33]	G1: Mailed educational materials (increase reach) G2: Individual academic detailing (increase ability) G3: Group academic detailing (increase ability)	Clinical outcomes (applicable for general public/patients) Rates of hospitalization (from electronic medical record)	Baseline and 1 year follow-up Objective	Baseline: 3692 Year 1: 2142	Baseline G1: 0.26, SD=0.94 G2: 0.26, SD=0.79 G3: 0.25, SD=0.77 Year 1 G1: 0.21, SD=0.79 G2: 0.18, SD=0.63 G3: 0.22, SD=0.69	Year 1 G1 vs. G2: 0.03* G1 vs. G3: 0.01* G2 vs. G3: 0.04*	Differences among individual patients Descriptive statistics (for determining M)

Table E-7. Key question 2 studies with a third outcome (continued)

Author, Year	Groups	Outcome #3, Exact Measure Used	Timing of Measurement, Data Source	N analyzed for This Outcome	Results by Group	Differences in Groups	Covariates Controlled for in Analysis, Statistical Methods Used
Soler et al., 2010[34]	G1: Control (not abstracted) G2: Training session on the SEPAR guidelines (increase ability) G3: G2 + portable-device for spirometry (multi-component)	Clinical outcomes (applicable for general public/patients) Differences in Treatment Regime (distribution of drugs prescribed according to the severity (mild vs. severe) of COPD before and after the training session. Treatment regimens: fixed combination, bronchodilators, corticoids and antibiotics.	Starting 90 days after training session Chart	G1: 1481, G2: 2119, G3: 5556	Long acting beta agonist: G1: 36.5% G2: 17.2% G3: 17% Anticholinergics: G1: 68.1% G2: 76.1% G3: 77.8% Theophylline G1: 5.5% G2:7.6% G3:4.9%	Long acting beta agonist: G2- G1: 19.3%, p=NR G3-G1: 19.5%, p=NR Anticholinergics: G2-G1: +8%, p= NR G3-G1: +9.7% Theophylline G2-G1: +2.1%, p=NR G3-G1: -0.6, p=NR	baseline value Logistic regression
Wolters et al., 2005[38]	G1: Control mailed guidelines (increase reach) G2: Intervention involving package for learning, supporting materials, decision tree, and information leaflets for patients (multi-component)	Behavior (applicable for clinicians) Adherence to guidelines. Number of patients referred to a urologist. The lower the referral rate, the better. More following of a watchful waiting policy	Up to 1 year post-intervention Prospective recording of patient data and management immediately after consultation with eligible patient	N = 187 G1: 92 G2: 95	Referral G1: 13, 14.5% G2: 2,2.1% Wait and see approach G1: 54, 58.7 G2: 61, 64.2%	Referral G1 vs. G2: 12.4%* OR:0.08 (0.02/0.40) Wait and see G1 vs. G2: 5.5%* OR:1.47 (0.66/3.28)	Age, group allocation, IPSS and BS Logistic regression

Abbreviations: ANOVA = ANalysis Of Variance; BS = Bother score; CaP = Cancer of the Prostate; CI = confidence interval; COPD = chronic obstructive pulmonary disease; CRC = colorectal cancer; DASH = Dietary Approaches to Stop Hypertension; dept = department; DVD = optical disc storage format; EMR = electronic medical record; FOBT = fecal occult blood test; G = group; ICS = inhaled corticosteroid; ICU = intensive care unit; IPSS = International Prostate Symptom Score; IQR = interquartile ratio; LHA = lay health advisor; LOS = length of stay; M = Mean; N = number; NR = not reported; NRS = Numeric rating scale; NS = not significant; OCP = oral contraceptive pill; PPD = puffs

per day; QOL = quality of life; RN = registered nurse; SD = standard deviation; SEPAR = Spanish Society of Pulmonology; TPV = tailored and targeted print and video; UC = usual care; WCB = Workers Compensation Board;

Table E-8. Key question 2 studies with a fourth outcome

Author, Year	Groups	Outcome #, Exact measure used	Timing of measurement, Data Source	N analyzed for this outcome	Results by group	Differences in Groups	Covariates controlled for in analysis, Statistical methods used,
Feldstein et al., 2006[13]	G1: Usual care (not abstracted) G2: EMR reminder (increase reach for clinicians) G3: EMR reminder and patient reminder (via letter with educational materials (multi-component)	Behavioral intentions to use or apply the evidence Physical activity (affirmative response to the query, "At least once a week, do you engage in any regular activity long enough to break a sweat?")	Baseline and 6 months after the intervention Patient self-report	G1: 33; G2: 41; G3: 42	Post intervention: Percent G1: Pre 21.2; Post 30.3 G2: Pre 22; Post 19.5 G3: Pre 26.2; Post 28.6	Difference: G1 vs. G2: -10.8;* G1 vs. G3: 1.7; * G2 vs. G3: -9.1* 95% CI: NR p=0.55 treatment and UC	Pre-survey response Regression analysis between treatment and usual care groups with the change of the post-survey response from pre-survey as the DV adjusting for pre-survey response

Table E-8. Key question 2 studies with a fourth outcome (continued)

Author, Year	Groups	Outcome #, Exact measure used	Timing of measurement, Data Source	N analyzed for this outcome	Results by group	Differences in Groups	Covariates controlled for in analysis, Statistical methods used,
Kennedy et al., 2003[19]	G1: Control (not abstracted) G2: Information (increase reach) G3: Interview (increase motivation)	Health-related decisions or behavior (applicable for general public/patients) Treatment undergone. For short-term, the treatment undergone is any treatment undergone up to 12 months. For long-term follow-up, the data presented are cumulative and refer to any treatment undergone during the period of the study.	The 6 and 12 month data were merged together to form a "short-term" follow-up dataset. 24 months is labeled "long-term" Self-report	Short-term Overall: 631 G2: 205 G3: 221 Long-term Overall: 729 G2: 232 G3: 253	Short-term G2: 170 (82.9%) G3: 186 (84.2%) Long-term G2: 204 (87.9%) G3: 212 (83.8%)	G2 vs. G3: 1.3%*, p=NR Long term G2 vs. G3: 4.1%*, p=NR	Consultant sex; Consultant year of qualification; Age; Baseline menorrhagia severity; Baseline knowledge; Previous treatment – D&C; Previous treatment – OCP; Previous treatment – hormonal drugs; Previous treatment – non-hormonal drugs; Duration of problem; Any previous surgery; Baseline preferences; Recruitment period; Length of follow-up logistic regression

Table E-8. Key question 2 studies with a fourth outcome (continued)

Author, Year	Groups	Outcome #, Exact measure used	Timing of measurement, Data Source	N analyzed for this outcome	Results by group	Differences in Groups	Covariates controlled for in analysis, Statistical methods used,
Marcus et al., 2009[25]	G1: Contact control treatment delayed group (not abstracted) G2: Telephone-based individualized feedback (increase motivation) G3: Print-based individualized feedback (increase reach)	Health-related decisions or behavior (applicable for general public/patients) Physical activity minutes per week via self-report on PAR (interviewer administered 7-day Physical Activity Recall) interview	Baseline, 6 months and 12 months Self-report	Overall N: 218 at 6 months; 205 at 12 months G1: 72 at 6 months, 69 at 12 months G2: 75 at 6 months, 70 at 12 months G3: 71 at 6 months, 66 at 12 months	G1: 6 months: 77.67 (SD=101.79); 12 months: 81.92 (SD=127.07) G2: 6 Months: 123.32 (SD=97.64) 12 Months: 100.59 (SD=119.68) G3: 6 Months: 129.49 (SD=156.46) 12 Months: 162.37 (SD=165.17)	Difference: 6 months: 6.17 12 months: 61.78 95% CI: NR P: 6 months: t = 0, p=0.8595 12 months: (t=2.72, p=0.0071)	Yes Analysis of covariance, adjusted for treatment effects for gender and seasonal differences. When overall test of between-groups differences was significant at the >05 level, the source of these differences was examined further using single-degree-of-freedom contrasts that compared the active treatment arms with each other as well as with the treatment delayed group.

Table E-8. Key question 2 studies with a fourth outcome (continued)

Author, Year	Groups	Outcome #, Exact measure used	Timing of measurement, Data Source	N analyzed for this outcome	Results by group	Differences in Groups	Covariates controlled for in analysis, Statistical methods used,
Murtaugh et al.,2005[27]	G1: Usual care (not abstracted) G2: Basic intervention e-mail reminder (increase reach) G3: Augmented intervention of e-mail reminder + package of supporting materials (multi-component)	Behavior (applicable for clinicians) % performing comprehensive CHF assessment	chart-review of subsequent RN visit, within 45 days of initial intake Chart	354	Overall N: 354 G1: 3.7% G2: 13.8% G3: 23.9%	Difference G2-G1: 10.1%, p=0.006 Difference G3-G1: 20.2%, P<0.001 Difference G3-G2: 10.1%*, CI and p=NR	Sociodemo-graphic variables of the RN (age, gender, race/ethnicity), Rn employment status, educational level and caseload; average baseline characteristics of patients cared for by each RN including health, functional status; geographic area where nurse provided care Predictive multivariate modeling
Partin et al., 2004[29]	G1: Usual care (not abstracted) G2: Pamphlet (increase reach) G3: Video (increase reach)	Health-related decisions or behavior (applicable for general public/patients) PSA testing	2 weeks and 1 year post-target appointment VA outpatient records	N = 893 G2: 295 G3: 308	Adjusted PSA rate w/in 2 weeks G2: 0.28 G3:0.29 Adjusted PSA w/in 1 year G2: 0.67 G3: 0.70	PSA w/in 2 weeks G2 vs. G3: -0.01*, p=NR PSA w/in 1 year G2 vs. G3: 0.70*, p=NR	None Logistic regression

Table E-8. Key question 2 studies with a fourth outcome (continued)

Author, Year	Groups	Outcome #, Exact measure used	Timing of measurement, Data Source	N analyzed for this outcome	Results by group	Differences in Groups	Covariates controlled for in analysis, Statistical methods used,
Rebbeck et al., 2006[31]	G1: Dissemination of guidelines by mail (increase reach) G2: Implementation group (multi-component)	Behavior (applicable for clinicians) Physiotherapist clinical practice - measured as the percentage of participating physiotherapists prescribing guideline recommendations	Before, during, and after the study self-report	Before G1: 13 G2: 14 After G1: 12 G2: 14 During G1: 12 G2: 14	% Reassure patient Before G1: 41% G2: 14% After G1: 18% G2: 57% During G1: 14% G2: 46% Advise to act as usual Before G1: 8 G2: 7 After G1: 18 G2: 67 During G1: 0 G2: 31 Prescribe function Before G1: 8 G2: 7 After G1: 0 G2: 25 During G1: 0 G2: 23	Reassure patient Difference (between G1 and G2 after trial): 39%* p=0.05 Advise to act as usual Difference (between G1 and G2 after trial): 49%* p=0.04 Prescribe function Difference (between G1 and G2 after trial): 25%* p=0.22 Prescribe exercise Difference (btn G1 and G2 after trial): 0%* p=1.00 Prescribe medication Difference (btn G1 and G2 after trial): 1%* p=0.10	NR Chi-square test

Table E-8. Key question 2 studies with a fourth outcome (continued)

Author, Year	Groups	Outcome #, Exact measure used	Timing of measurement, Data Source	N analyzed for this outcome	Results by group	Differences in Groups	Covariates controlled for in analysis, Statistical methods used,
Rebbeck et al., 2006[31] (continued)					Prescribe exercise Before G1: 92 G2: 100 After G1: 100 G2: 100 During G1: 100 G2: 83 Prescribe medication Before G1: 17 G2: 7 After G1: 9 G2: 8 During G1: 0 G2: 23		
Wolters et al., 2005[38]	G1: Control mailed guidelines (increase reach) G2: Intervention involving package for learning, supporting materials, decision tree, and information leaflets for patients (multi-component)	Discussions about the evidence Lifestyle advice given	Up to 1 years post baseline Self-report	N = 187 G1: 92 G2: 95	Lifestyle advice G1: 52, 56.5% G2: 58, 61.1%	Lifestyle advice G1 vs. G2: 4.6%* OR: 1.32 (0.48/3.63)	Age, group allocation, IPSS and BS Logistic regression analysis

Abbreviations: BS = Bother score; btn = between; CHF = congestive heart failure; CI = confidence interval; D&C = dilation and curettage; DV = dependent variable; EMR = electronic medical record; G = group; IPSS = International Prostate Symptom Score; N = number; NR = not reported; OCP = oral contraceptive pill; PAR = Stanford 7-Day Physical Activity Recall; PSA = prostate-specific antigen; RN = registered nurse; SD = standard deviation; vs. = versus; w/in = within.

Table E-9. Key question 2 studies with a fifth or sixth outcome

Author, Year	Groups	Outcome #, Exact Measure Used	Timing of Measurement, Data Source	N Analyzed for This Outcome	Results by Group	Differences in Groups	Covariates Controlled for in Analysis, Statistical Methods Used
Feldstein et al., 2006[13]	G1: Usual care (not abstracted) G2: EMR reminder (increase reach for clinicians) G3: EMR reminder and patient reminder (via letter with educational materials (multi-component)	Behavioral intentions to use or apply the evidence Total calcium intake mg/day	Baseline and 6 months after the intervention Patient self-report	G1: 22; G2: 33, G3: 37	Post intervention G1: Pre 1,308.6; Post 851.2 G2: Pre 1,116.5; Post 1311.4 G3: Pre 1,221.5; Post 1224.7	Difference: G1 vs. G2 -460.2;* G1 vs. G3 -373.5 * G2 vs. G3 -86.7* 95% CI: NR p=0.05 treatment and UC	Pre-survey response Same as Intermediate Outcome #1 but because of several high-intake outliers for total calcium intake, these data were Windsorized before further analysis

Table E-9. Key question 2 studies with a fifth or sixth outcome (continued)

Author, Year	Groups	Outcome #, Exact Measure Used	Timing of Measurement, Data Source	N Analyzed for This Outcome	Results by Group	Differences in Groups	Covariates Controlled for in Analysis, Statistical Methods Used
Marcus et al., 2009[25]	G1: Contact control treatment delayed group (not abstracted) G2: Telephone-based individualized feedback (increase motivation) G3: Print-based individualized feedback (increase reach)	Behavior (applicable for clinicians) Achieving 150 minutes of physical activity per week	Baseline, 6 months and 12 months Self-report	Overall N: 218 at 6 months; 205 at 12 months G1: 72 at 6 months, 69 at 12 months G2: 75 at 6 months, 70 at 12 months G3: 71 at 6 months, 66 at 12 months	NR	OR and 95% CI Difference: G2 compared to G1: 6 Months: OR,3.30 [1.66, 7.22]; 12 Months: OR, 1.50 [0.67, 3.33]. G3 compared to G1: 6 Months: OR,2.95 [1.41, 6.19]; 12 Months OR,5.31 [2.47, 11.39]. G3 compared to G2 6 Months: OR,1.18 [0.62, 2.24]; 12 Months: OR,3.55 [1.76, 7.16]	Yes F tests for between group differences and logistics regression for binary outcomes
		Behavior (applicable for clinicians) Validation of the PAR with the Actigraph: comparison of activity above and below the Actigraph count threshold considered to be moderate physical activity	6 and 12 months Self-report and objective measurement via Actigraph	NR - sub-sample but N not reported	G2: NR G3: NR	Difference: 6 Months: Pearson correlation of .30, p=0.0224; 12 Months Pearson correlation of .32, p=0.019	Behavior (applicable for clinicians) validation of the PAR with the Actigraph: comparison of activity above and below the Actigraph count threshold considered to be moderate physical activity

Table E-9. Key question 2 studies with a fifth or sixth outcome (continued)

Author, Year	Groups	Outcome #, Exact Measure Used	Timing of Measurement, Data Source	N Analyzed for This Outcome	Results by Group	Differences in Groups	Covariates Controlled for in Analysis, Statistical Methods Used
Murtaugh et al.,2005[27]	G1: Usual care (not abstracted) G2: Basic intervention e-mail reminder (increase reach) G3: Augmented intervention of e-mail reminder + package of supporting materials (multi-component)	Behavior (applicable for clinicians) % assessing patient diet	Chart-review of subsequent RN visit, within 45 days of Initial intake Chart	354	Overall N: 354 G1: 27.6% G2: 38.2% G3: 48.7%	Difference G2-G1: 10.6%, p=0.076 Difference G3-G1: 21.1%, p=0.001 Difference G3-G2: 10.5%*, CI and p=NR	Socio-demographic variables of the RN (age, gender, race/ethnicity), Rn employment status, educational level and caseload; average baseline characteristics of patients cared for by each RN including health, functional status; geographic area where nurse provided care Predictive multivariate modeling

Table E-9. Key question 2 studies with a fifth or sixth outcome (continued)

Author, Year	Groups	Outcome #, Exact Measure Used	Timing of Measurement, Data Source	N Analyzed for This Outcome	Results by Group	Differences in Groups	Covariates Controlled for in Analysis, Statistical Methods Used
Murtaugh et al.,2005[27] (continued)		Behavior (applicable for clinicians) % assessing medication side-effects effects	Chart-review of subsequent RN visit, within 45 days of initial intake Chart	354	Overall N: 354 G1: 12.7% G2: 15.3% G3: 23.6%	Difference G2-G1: 2.6%, p=0.558 Difference G3-G1: 10.9%, p=0.030 Difference G3-G2: 8.3%*, CI and p= NR	Socio-demographic variables of the RN (age, gender, race/ethnicity), Rn employment status, educational level and caseload; average baseline characteristics of patients cared for by each RN including health, functional status; geographic area where nurse provided care Predictive multivariate modeling

Abbreviations: CI = confidence interval; EMR = electronic medical record; mg = milligram; NR = not reported; OR = odds ratio; PAR = Stanford 7-Day Physical Activity Recall; RN = registered nurse; vs. = versus.

Table E-10. Key question 2 other analysis information

Author, Year	Groups	Analysis Adjusted for Multiple Comparisons, if Applicable	Analysis Adjusted for Clustering Effect, if Applicable	Intention to Treat Analysis, if Applicable	Additional Outcomes	Other Notes
Bahrami et al., 2004[1]	G1: Mailed guideline (increase reach) G2: Guideline + AF (not abstracted) G3: CAL (increase ability) G4: CAL + AF (not abstracted)	NR	Yes	Yes	NR	
Banait et al., 2003[2]	G1: Mailed guidelines (increase reach) G2: Educational outreach (Multi-component)	NR	NA	Yes	Prescription costs for acid-suppressing drugs.	
Beaulieu et al., 2004[3]	G1: Control (not abstracted) G2: Guideline (increase reach) G3: Guideline + reminder notice and stickers for patients' charts (multi-component)	NR	Yes	No	NR	
Becker et al., 2008[4]	G1: Mailed guideline (Increase clinician reach) G2: Guideline implementation (multi-component, clinicians only) G3: Guideline implementation and motivational counseling directed at patient (multi-component, clinicians and patients)	NR	Yes, accounted for effects of clustering	Yes,	Days in pain^ Days of sick leave	

Table E-10. Key question 2 other analysis information (continued)

Author, Year	Groups	Analysis Adjusted for Multiple Comparisons, if Applicable	Analysis Adjusted for Clustering Effect, if Applicable	Intention to Treat Analysis, if Applicable	Additional Outcomes	Other Notes
Bekkering et al., 2005[5]	G1: Received guidelines by mail (increase reach) G2: Received guidelines + active training strategy (multi-component)	NR	Yes	Yes	Sick leave (number of days off work in the last 6 weeks); pain coping strategies; pain beliefs	
Bishop and Wing, 2006[6]	G1: Control (not abstracted) G2: Physician only (increase reach) G3: Physician and patient (multi-component)	NR	NA	NR		The authors only compared G2 and G3 with the no treatment control group. They did not see if the groups were different from each other. Also, they do not provide any figures, tables, or data for the >12 week measures. They just say that there was no change seem in the recommended use of ongoing supervised exercise programs. It seems like all p-values apply to comparisons with the control group, not among G2 and G3. Bedrest data are for 5-12 weeks, while other data are for 0-4 weeks.

Table E-10. Key question 2 other analysis information (continued)

Author, Year	Groups	Analysis Adjusted for Multiple Comparisons, if Applicable	Analysis Adjusted for Clustering Effect, if Applicable	Intention to Treat Analysis, if Applicable	Additional Outcomes	Other Notes
Campbell et al., 2004[7]	G1: Control (not abstracted) G2: LHA (increase motivation) G3: TPV (multi-component) G4: TPV and LHA (multi-component)	Yes	Yes	NR	Calories from fat Process measures	
Carney et al., 2005[8]	G1: Mailed health information (increase reach) G2: Telephone counseling (increase motivation)	NA	NA	NA	Stages of change (for G2 only)	
Christakis et al., 2006[9]	G1: Usual care (not abstracted) G2: Parental content Alone (increase reach) G3: Provider notification alone (not abstracted) G4: Parental content and provider notification (multi-component)	NR	results were adjusted for clustering with respect to physician	Yes	NR	Occurred in only 1 practice, 40% of those eligible did not enroll, some possible contamination
Davis et al., 2004[10]	G1: Control - guidelines by mail (increase reach) G2: Intermediate (multi-component) G3: High intervention (multi-component)	NR	Yes	Yes	Process of care outcomes (although these were primarily assessed in G3) Nature of seizures Perceived severity of seizures Perceived adverse drug effects	

Table E-10. Key question 2 other analysis information (continued)

Author, Year	Groups	Analysis Adjusted for Multiple Comparisons, if Applicable	Analysis Adjusted for Clustering Effect, if Applicable	Intention to Treat Analysis, if Applicable	Additional Outcomes	Other Notes
Eaton et al., 2011[11]	G1: 1-hour academic detailing (increase clinician ability) G2: Academic detailing plus a patient education toolkit, a computer kiosk with patient activation software, and a PDA-based decision support tool (multi-component)	No	Yes	Yes	NR	
Elder et al., 2005[12]	G1: Culturally targeted print-materials + activity inserts (increase reach) G2: Tailored print materials + activity inserts + supporting materials (multi-component). G3: Tailored print materials + in-person promotora (multi-component)	Yes	NA	No	BMI Energy^ Total fat^ Total saturated fat ^ Soluable dietary fiber Insoluable dietary fiber Total carbs^ Glucose^ Fructose^ Sucrose^	

Table E-10. Key question 2 other analysis information (continued)

Author, Year	Groups	Analysis Adjusted for Multiple Comparisons, if Applicable	Analysis Adjusted for Clustering Effect, if Applicable	Intention to Treat Analysis, if Applicable	Additional Outcomes	Other Notes
Feldstein et al., 2006[13]	G1: Usual care (not abstracted) G2: EMR reminder (increase reach for clinicians) G3: EMR reminder and patient reminder (via letter with educational materials (multi-component)	NR	NA	NR	rPatient Satisfaction Participants asked to rate the overall care they received for their bones on a scale of 1 (terrible) to 5 (superb). Mean change: G1: -0.07; G2: 0.07 and G3: 0.08. Differences not significant. p=0.81.	
Gattellari et al., 2005[14]	G1: Leaflet (increase reach) G2: Video (increase reach) G3: Booklet (increase reach)					
Hagmolen et al., 2008[15]	G1: Guideline dissemination (increase reach) G2: Guideline dissemination + educational program (increase ability) G3: Guideline dissemination + educational program + individualized treatment advice based on airway responsiveness and symptoms (multi-component)	NR	Yes, accounted for effects of clustering	Yes, results analyzed on an intention to treat basis	Number of symptom free days FEV1 % of predicted PEF variability % β2-agonist, ppd (GP, 1 yr)	

Table E-10. Key question 2 other analysis information (continued)

Author, Year	Groups	Analysis Adjusted for Multiple Comparisons, if Applicable	Analysis Adjusted for Clustering Effect, if Applicable	Intention to Treat Analysis, if Applicable	Additional Outcomes	Other Notes
Jain et al., 2006[16]	G1: Passive intervention- guidelines by mail (increase reach) G2: Active intervention (multi-component)	No	YES	NA	Site's use of feeding protocol Head of bed elevation EN initied within 48 hours Use of glutamine Use of motility agents in EN patients Small bowel feeding in patients with feed interrupted due to high gastric residuals Lipid use in PN patients	
Jousimaa et al., 2002[17]	G1: Computerized version of guidelines (increase ability) G2: Textbook-based version of guidelines (increase reach)	No	Yes	No	None	
Junghans et al., 2007[18]	G1: Conventional guideline (increase reach) G2: Ratings about specific patients in vignettes (increase motivation)	NR	Yes	NA		

Table E-10. Key question 2 other analysis information (continued)

Author, Year	Groups	Analysis Adjusted for Multiple Comparisons, if Applicable	Analysis Adjusted for Clustering Effect, if Applicable	Intention to Treat Analysis, if Applicable	Additional Outcomes	Other Notes
Kennedy et al., 2003[19]	G1: Control (not abstracted) G2: Information (increase reach) G3: Interview (increase motivation)	NR	Yes	Yes	Anxiety Agreement between preferences and treatments undergone EQ-5D instrument. Severity of menorrhagia Patient satisfaction Cost-effectiveness	

Table E-10. Key question 2 other analysis information (continued)

Author, Year	Groups	Analysis Adjusted for Multiple Comparisons, if Applicable	Analysis Adjusted for Clustering Effect, if Applicable	Intention to Treat Analysis, if Applicable	Additional Outcomes	Other Notes
King et al., 2007[20]	G1: Attention control (not abstracted) G2: Counselor via phone (increase motivation) G3: Automated counselor via phone (increase reach)	No	NA	Yes	Self-rated functioning physical fitness psychological well-being Quality of life	At 6 and 12 months, both G2 & G3 had greater PAR & CHAMPS mean energy expenditure than G1, p=0.01. At 6 and 12 months, both G2 & G3 had greater PAR & CHAMPS mean minutes/week than G1, p=0.01. At 6 months, both G2 & G3 had greater # of PAR & CHAMPS days/week of MOD+ activity than G1, p=0.001 for G2 vs. G1 and p=0.01 for G3 vs. G1 For the objective physical activity measure, the combined intervention arms showed more physical activity relative to G1.

Table E-10. Key question 2 other analysis information (continued)

Author, Year	Groups	Analysis Adjusted for Multiple Comparisons, if Applicable	Analysis Adjusted for Clustering Effect, if Applicable	Intention to Treat Analysis, if Applicable	Additional Outcomes	Other Notes
Laprise et al., 2009[21]	G1: CME (increase ability) G2: CME + practice enablers and reinforcers (multi-component)	NR	Yes	NA	RXN of a pharmacological agent, referral, or counseling in smokers*	
Lien et al., 2007[22], Svetkey et al., 2003[23], Young et al., 2009[24]	G1: Advice only (increase reach) G2: Advice + behavioral counseling using established intervention (multi-component) G3: Established intervention + DASH dietary recommendations (multi-component)	Yes	NA	YES	Physical activity Fitness ^ Alcohol intake Urine collections^ Dietary recalls^ BUN Phosphorus Fruits/vegetables, dairy Calcium Total fat^ Saturated fat^	
Marcus et al., 2009[25]	G1: Contact control treatment delayed group (not abstracted) G2: Telephone-based individualized feedback (increase motivation) G3: Print-based individualized feedback (increase reach)	NA	NA	Yes	Cognitive processes Measured via self-reports Differences between intervention arms at 6 and 12 months were not significant.	
Maxwell et al.,2010[26]	G1: Control (not abstracted) G2: Educational session + letter to provider (multi-component) G3: Educational session + letter to provider + FOBT kit (multi-component)					

Table E-10. Key question 2 other analysis information (continued)

Author, Year	Groups	Analysis Adjusted for Multiple Comparisons, if Applicable	Analysis Adjusted for Clustering Effect, if Applicable	Intention to Treat Analysis, if Applicable	Additional Outcomes	Other Notes
Murtaugh et al.,2005[27]	G1: Usual care (not abstracted) G2: Basic intervention e-mail reminder (increase reach) G3: Augmented intervention of e-mail reminder + package of supporting materials (multi-component)	NR	NA	NR	Also looked at % recording medication knowledge and medication adherence; also looked at % recording self management instructions (weighing self, managing weight gain, low salt diet, managing meds, improving adherence, contacting MD--some of these were positive/others negative).	None of the methods to assess the outcome measures or measure details are included
Paradis et al.,2011[28]	G1: Paper handouts (increase reach) G2: Educational DVD (increase reach)	NR	NA	Yes	Parent satisfaction with enrollment visit; staff and provider satisfaction with DVD	
Partin et al., 2004[29]	G1: Usual care (not abstracted) G2: Pamphlet (increase reach) G3: Video (increase reach)	No	NA	No, but they conducted analyses with only those that reported exposure to the interventions. See Tab 4 Column AL	NA	

Table E-10. Key question 2 other analysis information (continued)

Author, Year	Groups	Analysis Adjusted for Multiple Comparisons, if Applicable	Analysis Adjusted for Clustering Effect, if Applicable	Intention to Treat Analysis, if Applicable	Additional Outcomes	Other Notes
Rahme et al., 2005[30]	G1: No treatment control (not abstracted) G2: Decision tree (increase ability) G3: Workshop (increase ability) G4: Workshop + decision tree (multi-component)	NR	Yes	Yes		
Rebbeck et al., 2006[31]	G1: Dissemination of guidelines by mail (increase reach) G2: Implementation group (multi-component)	No	Yes	No	Change in symptoms, patient satisfaction, physiotherapist satisfaction, cost of care	
Rimer et al., 2001[32]	G1: No treatment control/usual care (not abstracted) G2: Tailored print (increase reach) G3: Tailored print + telephone counselling (multi-component)					TP was different than TP +TC (G2 v G3) for 12-month rates but this disappeared by 15 months, so I felt we didn't need that; clinical breast exam differences were NS; didn't control for baseline starting point.
Simon et al., 2005[33]	G1: Mailed educational materials (increase reach) G2: Individual academic detailing (increase ability) G3: Group academic detailing (increase ability)	No	Yes	Yes	Outpatient visits, average per-person cost of antihypertensive medications	none

Table E-10. Key question 2 other analysis information (continued)

Author, Year	Groups	Analysis Adjusted for Multiple Comparisons, if Applicable	Analysis Adjusted for Clustering Effect, if Applicable	Intention to Treat Analysis, if Applicable	Additional Outcomes	Other Notes
Soler et al., 2010[34]	G1: Control (not abstracted) G2: Training session on the SEPAR guidelines (increase ability) G3: G2 + portable-device for spirometry (multi-component)	No	NA	yes	fixed combination treatment, short acting beta agonists, corticosteroid, antibiotic use	
Sullivan et al., 2010[35]	G1: VA guidelines (increase reach) G2: COPE: web-based education program (increase ability)	NR	NA	Yes	Self-rated competence; physician satisfaction in caring for patients with CNCP; physician patient-centeredness; satisfaction with training (for those assigned to COPE)	
Watson et al., 2002[36]	G1: Guideline materials by postal mail (increase reach) G2: EO session and guidelines (increase ability) G3: CPE session and guidelines (increase ability) G4: Guidelines + EO and CPE (multi-component)	Yes	Yes	NR		

Table E-10. Key question 2 other analysis information (continued)

Author, Year	Groups	Analysis Adjusted for Multiple Comparisons, if Applicable	Analysis Adjusted for Clustering Effect, if Applicable	Intention to Treat Analysis, if Applicable	Additional Outcomes	Other Notes
Wetter et al., 2006[37]	G1: Single standard telephone-counseling session (increase reach) G2: Multiple enhanced telephone counseling sessions (multi-component)	No	NA	No	NA	
Wolters et al., 2005[38]	G1: Control mailed guidelines (increase reach) G2: Intervention involving package for learning, supporting materials, decision tree, and information leaflets for patients (multi-component)	NR	Yes	No	History taking Diagnostic procedures Consultation lasting more than 15 min^ Use of Wait and See approach	A non-response analysis was done to see if the GPs that did not complete the educational program or recruit patients were any different than those that did. No differences were found.
Wright et al., 2008[39]	G1: Standardized lecture by expert opinion leader (increase motivation) G2: Standardized lecture by expert opinion leader + academic detailing and a toolkit (multi-component)	No	Yes	No	NA	

Abbreviations: AF = audit and feedback; BMI = body mass index; BUN = blood urea nitrogen; CAL = computer-assisted learning; CHAMPS = Community Healthy Activities Model Program for Seniors; CME = continuing medical education; CNCP = Chronic non-cancer pain; COPE = Compassionate Options for Progressive Eldercare; DASH = Dietary Approaches to Stop Hypertension; DVD = optical disc storage format; EMR = electronic medical record; EN = enteral nutrition; EQ-5D = EuroQol-5 dimensions; FEV% = Forced Percentual Expiratory Volume; FOBT = fecal occult blood test; G = group; GP = general practitioner; LHA = lay health advisor; MD = medical doctor; MOD = moderate intensity or more vigorous; NA = not applicable; NR = not reported; PAR = Stanford 7-Day Physical Activity Recall; PDA = personal digital assistant; PPD = puffs per day; RXN = reaction; SEPAR = Spanish Society of Pulmonology; TP = Tailored print; TC = Tailored print and telephone counseling; TPV = tailored and targeted print and video; vs. = versus

References

1. Bahrami M, Deery C, Clarkson JE, et al. Effectiveness of strategies to disseminate and implement clinical guidelines for the management of impacted and unerupted third molars in primary dental care, a cluster randomised controlled trial. Br Dent J. 2004 Dec 11;197(11):691-6; discussion 88. PMID: 15592551.

2. Banait G, Sibbald B, Thompson D, et al. Modifying dyspepsia management in primary care: a cluster randomised controlled trial of educational outreach compared with passive guideline dissemination. Br J Gen Pract. 2003 Feb;53(487):94-100. PMID: 12817353.

3. Beaulieu MD, Brophy J, Jacques A, et al. Drug treatment of stable angina pectoris and mass dissemination of therapeutic guidelines: a randomized controlled trial. QJM. 2004 Jan;97(1):21-31. PMID: 14702508.

4. Becker A, Leonhardt C, Kochen MM, et al. Effects of two guideline implementation strategies on patient outcomes in primary care: a cluster randomized controlled trial. Spine (Phila Pa 1976). 2008(5):473-80. PMID: 18317189.

5. Bekkering GE, Hendriks HJ, van Tulder MW, et al. Effect on the process of care of an active strategy to implement clinical guidelines on physiotherapy for low back pain: a cluster randomised controlled trial. Qual Saf Health Care. 2005(2):107-12. PMID: 15805455.

6. Bishop PB, Wing PC. Knowledge transfer in family physicians managing patients with acute low back pain: a prospective randomized control trial. Spine J. 2006 May-Jun;6(3):282-8. PMID: 16651222.

7. Campbell MK, James A, Hudson MA, et al. Improving multiple behaviors for colorectal cancer prevention among African American church members. Health Psychol. 2004;23(5):492-502. PMID: 15367069.

8. Carney PA, Harwood BG, Greene MA, et al. Impact of a telephone counseling intervention on transitions in stage of change and adherence to interval mammography screening (United States). Cancer Causes Control. 2005 Sep;16(7):799-807. PMID: 16132790.

9. Christakis DA, Zimmerman FJ, Rivara FP, et al. Improving pediatric prevention via the internet: a randomized, controlled trial. Pediatrics. 2006 Sep;118(3):1157-66. PMID: 16951011.

10. Davis J, Roberts R, Davidson DL, et al. Implementation strategies for a Scottish national epilepsy guideline in primary care: results of the Tayside Implementation of Guidelines in Epilepsy Randomized (TIGER) trial. Epilepsia. 2004 Jan;45(1):28-34. PMID: 14692904.

11. Eaton CB, Parker DR, Borkan J, et al. Translating cholesterol guidelines into primary care practice: a multimodal cluster randomized trial. Ann Fam Med. 2011 Nov-Dec;9(6):528-37. PMID: 22084264.

12. Elder JP, Ayala GX, Campbell NR, et al. Long-term effects of a communication intervention for Spanish-dominant Latinas. Am J Prev Med. 2006 Aug;31(2):159-66. PMID: 16829333.

13. Feldstein A, Elmer PJ, Smith DH, et al. Electronic medical record reminder improves osteoporosis management after a fracture: a randomized, controlled trial. J Am Geriatr Soc. 2006 Mar;54(3):450-7. PMID: 16551312.

14. Gattellari M, Ward JE. A community-based randomised controlled trial of three different educational resources for men about prostate cancer screening. Patient Educ Couns. 2005;57(2):162-82. PMID: 15911190.

15. Hagmolen of ten Have W, van den Berg NJ, van der Palen J, et al. Implementation of an asthma guideline for the management of childhood asthma in general practice: a randomised controlled trial. Prim Care Respir J. 2008 Jun;17(2):90-6. PMID: 18274692.

16. Jain MK, Heyland D, Dhaliwal R, et al. Dissemination of the Canadian clinical practice guidelines for nutrition support: results of a cluster randomized controlled trial. Crit Care Med. 2006 Sep;34(9):2362-9. PMID: 16850001.

17. Jousimaa J, Mäkelä M, Kunnamo I, et al. Primary care guidelines on consultation practices: the effectiveness of computerized versus paper-based versions. A cluster randomized controlled trial among newly qualified primary care physicians. Int J Technol Assess Health Care. 2002 Sum(3):586-96. PMID: 12391951.

18. Junghans C, Feder G, Timmis AD, et al. Effect of patient-specific ratings vs conventional guidelines on investigation decisions in angina: Appropriateness of Referral and Investigation in Angina (ARIA) Trial. Arch Intern Med. 2007 Jan 22;167(2):195-202. PMID: 17242322.

19. Kennedy ADM, Sculpher MJ, Coulter A, et al. A multicentre randomised controlled trial assessing the costs and benefits of using structured information and analysis of women's preferences in the management of menorrhagia. Health Technol Assess. 2003;7(8):1-76. PMID: 12633525.

20. King AC, Friedman R, Marcus B, et al. Ongoing physical activity advice by humans versus computers: The Community Health Advice by Telephone (CHAT) trial. Health Psychol. 2007 Nov;26(6):718-27. PMID: 18020844.

21. Laprise R, Thivierge R, Gosselin G, et al. Improved cardiovascular prevention using best CME practices: a randomized trial. J Contin Educ Health Prof. 2009 Winter;29(1):16-31. PMID: 19288563.

22. Lien LF, Brown AJ, Ard JD, et al. Effects of PREMIER lifestyle modifications on participants with and without the metabolic syndrome. Hypertension. 2007 Oct;50(4):609-16. PMID: 17698724.

23. Svetkey LP, Harsha DW, Vollmer WM, et al. Premier: A clinical trial of comprehensive lifestyle modification for blood pressure control: rationale, design and baseline characteristics. Ann Epidemiol. 2003 Jul;13(6):462-71. PMID: 12875806.

24. Young DR, Vollmer WM, King AC, et al. Can individuals meet multiple physical activity and dietary behavior goals? Am J Health Behav. 2009 May-Jun;33(3):277-86. PMID: 19063649.

25. Marcus BH, Napolitano MA, King AC, et al. Telephone versus print delivery of an individualized motivationally tailored physical activity intervention: Project STRIDE. Health Psychol. 2007 Jul;26(4):401-9. PMID: 17605559.

26. Maxwell AE, Bastani R, Danao LL, et al. Results of a community-based randomized trial to increase colorectal cancer screening among Filipino Americans. Am J Public Health. 2010 Nov;100(11):2228-34. PMID: 20864724.

27. Murtaugh CM, Pezzin LE, McDonald MV, et al. Just-in-time evidence-based e-mail "reminders" in home health care: impact on nurse practices. Health Serv Res. 2005 Jun;40(3):849-64. PMID: 15960694.

28. Paradis HA, Conn KM, Gewirtz JR, et al. Innovative delivery of newborn anticipatory guidance: a randomized, controlled trial incorporating media-based learning into primary care. Acad Pediatr. 2011 Jan-Feb;11(1):27-33. PMID: 21272821.

29. Partin MR, Nelson D, Radosevich D, et al. Randomized trial examining the effect of two prostate cancer screening educational interventions on patient knowledge, preferences, and behaviors. J Gen Intern Med. 2004 Aug;19(8):835-42. PMID: 15242468.

30. Rahme E, Choquette D, Beaulieu M, et al. Impact of a general practitioner educational intervention on osteoarthritis treatment in an elderly population. Am J Med. 2005 Nov;118(11):1262-70. PMID: 16271911.

31. Rebbeck T, Maher CG, Refshauge KM. Evaluating two implementation strategies for whiplash guidelines in physiotherapy: a cluster randomised trial. Aust J Physiother. 2006;52(3):165-74. PMID: 16942451.

32. Rimer BK, Halabi S, Sugg Skinner C, et al. The short-term impact of tailored mammography decision-making interventions. Patient Educ Couns. 2001 Jun;43(3):269-85. PMID: 11384825.

33. Simon SR, Majumdar SR, Prosser LA, et al. Group versus individual academic detailing to improve the use of antihypertensive medications in primary care: a cluster-randomized controlled trial. Am J Med. 2005 May;118(5):521-8. PMID: 15866255.

34. Soler N, Ballester E, Martin A, et al. Changes in management of chronic obstructive pulmonary disease (COPD) in primary care: EMMEPOC study. Respir Med. 2010 Jan;104(1):67-75. PMID: 20122630.

35. Sullivan MD, Gaster B, Russo J, et al. Randomized trial of web-based training about opioid therapy for chronic pain. Clin J Pain. 2010 Jul;26(6):512-7. PMID: 20551726.

36. Watson MC, Bond CM, Grimshaw JM, et al. Educational strategies to promote evidence-based community pharmacy practice: a cluster randomized controlled trial (RCT). Fam Pract. 2002 Oct;19(5):529-36. PMID: 12356708.

37. Wetter DW, Mazas C, Daza P, et al. Reaching and treating Spanish-speaking smokers through the National Cancer Institute's Cancer Information Service. A randomized controlled trial. Cancer. 2007 Jan 15;109(2 Suppl):406-13. PMID: 17149758.

38. Wolters R, Wensing M, Klomp M, et al. Effects of distance learning on clinical management of LUTS in primary care: a randomised trial. Patient Educ Couns. 2005 Nov;59(2):212-8. PMID: 16257627.

39. Wright FC, Gagliardi AR, Law CH, et al. A randomized controlled trial to improve lymph node assessment in stage II colon cancer. Arch Surg. 2008 Nov;143(11):1050-5; discussion 5. PMID: 19015462.

40. Bekkering GE, van Tulder MW, Hendriks EJ, et al. Implementation of clinical guidelines on physical therapy for patients with low back pain: randomized trial comparing patient outcomes after a standard and active implementation strategy. Phys Ther. 2005(6):544-55. PMID: 15921475.

Appendix F. Evidence Tables for Key Question 3

Table F-1. Key question 3 study design details

Author, Year	Research objective	Funding Source	Geographic Location, Setting Type, Setting Description	Study Design	Primary Outcomes	Measurement Intervals	Other Notes
Akl et al., 2012[1]	To compare different wording approaches for conveying the strength of health care recommendations.	Other	United States and Canada Academic health care institutions Medical residency program large group teaching sessions	fRCT	Appropriate or inappropriate course of action	Immediate posttest	Study was not funded
Brewer et al., 2012[2]	Conducted an experiment with early stage breast cancer patients that compared risk communication formats of varying complexity that used elements from the Oncotype DX report.	Government	USA Academic health care institutions University of NC Breast Clinic	Quasi-Experimental	Accuracy of Risk Perception (gist), % incorrect/error Accuracy of Risk Perception (verbatim), % incorrect/error Attitude toward the test result	Immediate posttest	
Han et al., 2011[3] (Experiment 1)	To explore the effect of communicating uncertainty on people's responses to comparative risk information.	Government	United States Other NA	fRCT	Risk perception Worry	Immediate posttest.	Web-based
Han et al., 2011[3] (Experiment 2)	To explore the effect of novel visual and textual representations of uncertainty.	Government	United States Other NA	Randomized trial	Risk perception Worry	Immediate posttest	Web-based

Table F-1. Key question 3 study design details (continued)

Author, Year	Research objective	Funding Source	Geographic Location, Setting Type, Setting Description	Study Design	Primary Outcomes	Measurement Intervals	Other Notes
McCormack et al., 2011[4]	To examine the effects of a community-based intervention on decisions about PSA screening using multiple measures of IDM.	Government	USA Community-based settings Intervention groups were 2 NC communities and their community-based organizations (senior, faith-based, fraternal, fitness, and recreationals), control was a 3rd NC community	Non-randomized trial	Prostate CA screening and treatment knowledge Self-efficacy PSA screening decision Preferred level of involvement Belief that screening is a decision	Baseline, 6 months, and 12 months	
Perneger et al., 2010[5] and 2011 [6]	To examine whether information about risks and benefits of cancer screening leads to higher test refusal rates and satisfaction with the decision that was made.	Foundation or non-profit	Switzerland Community-based settings Mailed survey to adults living in the Swiss canton of Geneva	fRCT	refusal rates for screening composite decision evaluation score	Baseline	
Schwartz et al., 2011[7]	To assess the US public's understanding of the meaning of FDA drug approval and test how brief explanations communicating drug uncertainties affect consumer choice	Foundation or non-profit	USA Community-based settings Nationally representative sample of Americans recruited from a research panel of approximately 30,000 households	Randomized trial	Choice of the better drug, either the drug that affects more distal outcomes or the one that's been on the market the longest	Immediately following intervention	Study part of a larger study and the larger study oversampled minorities

Table F-1. Key question 3 study design details (continued)

Author, Year	Research objective	Funding Source	Geographic Location, Setting Type, Setting Description	Study Design	Primary Outcomes	Measurement Intervals	Other Notes
Steginga et al., 2002[8]	The present study investigated the effect of varying situational risk and ambiguity on desire for involvement in a hypothetical scenario about treatment for prostate cancer and examined the nature of the psychological predictors of desire for involvement in decision making about the disease.	Unspecified	Australia Community-based settings Urban and rural Queensland	Randomized trial	Involvement in decision making	Immediate post-test	

Abbreviations: CA = cancer; DX = diagnosis; FDA = Food and Drug Administration; fRCT = factorial randomized controlled trial; IDM = informed decision making; NA = not applicable; NC = North Carolina; PSA = prostate-specific antigen; USA = United States of America

Table F-2. Key question 3 sample characteristics, part 1

Author, Year	Groups	Sampling Strategy, Unit of Randomization, Process of Randomization	Inclusion/ Exclusion Criteria	N Eligible	N Randomized	N Com-pleters	N Analyzed	Other Notes
Akl et al., 2012[1]	G1: Strong and weak wording for or against guideline-supported behavior (ACCP): • "we recommend" • "we suggest" • "we suggest...not" "we recommend…not" G2: Strong and weak wording for or against guideline-supported behavior (NICE): • "clinicians should" • "clinicians might" • "clinicians might not" • "clinicians should not" G3: Strong and weak recommendations for or against guideline-supported behavior (GRADE): • "we recommend" • "we conditionally recommend" • "we conditionally recommend…not" • "we recommend…not"	Convenience Individual NR	Inclusion: residents in four Internal Medicine and two Family Medicine residency programs in five universities in Canada and the United States	N = 441	N = 341 G1: 114 G2: 111 G3: 118	N = 341 G1: 114 G2: 111 G3: 118	N = 331 G1: 114 G2: 111 G3: 118	

Table F-2. Key question 3 sample characteristics, part 1 (continued)

Author, Year	Groups	Sampling Strategy, Unit of Randomization, Process of Randomization	Inclusion/ Exclusion Criteria	N Eligible	N Randomized	N Com-pleters	N Analyzed	Other Notes
Brewer et al., 2012[2]	G1: Percent + verbal descriptor (reference) G2: G1 + risk continuum graphic (reference) G3: G2 + confidence interval (precision) G4: G3 + risk score + graph (NA) G5: Oncotype DX report (precision) G5: Icon array (reference) Note: Each format presents the likelihood of recurrence of breast cancer	Convenience NA NA However, vignette order randomized and components of vignettes randomized as follows. Participants viewed 6 of the 36 possible vignettes. Authors counterbalanced the vignettes such that they randomly assigned each vignette to one of the 6 risk magnitudes without replacement and to one of the six risk formats using a Latin square.	Inclusion: Women being treated for early stage breast cancer at the University of North Carolina Breast Clinic (Chapel Hill, NC) between May 2009 and November 2010, who were eligible to receive genomic recurrence risk testing by Oncotype DX (diagnosed with stages I or II, node-negative, hormone receptor–positive breast cancer), whether or not they actually received the testing. Exclusion: women who were non-English speaking, incarcerated, had a second primary cancer diagnosis or other life threatening co-morbid disease, or had a history of a serious psychiatric diagnosis.	N= 225	N = 143	N = 133	N = 133	

Table F-2. Key question 3 sample characteristics, part 1 (continued)

Author, Year	Groups	Sampling Strategy, Unit of Randomization, Process of Randomization	Inclusion/ Exclusion Criteria	N Eligible	N Randomized	N Com-pleters	N Analyzed	Other Notes
Han et al., 2011[3] (Experiment 1)	G1: Point estimate in text G2: Point estimate as graph G3: Range in text G4: Range as graph NOTE: Each format tested before and after delivery of information about the populations average risk of colon cancer (6% vs. 2 to 10%).	Convenience Individual NR	Eligibility criteria for the current study included age ≥ 40 years and no prior history of colon cancer.	N = 240	N = 240 G1: 60 G2: 60 G3: 60 G4: 60	N = 240 G1: 60 G2: 60 G3: 60 G4: 60	N = 240 G1: 60 G2: 60 G3: 60 G4: 60	
Han et al., 2011[3] (Experiment 2)	G1: Range in text (precision) G2: Range in text + solid bar graph (precision) G3: Range in text + blurred bar graph (precision)	Convenience Individual NR	Eligibility criteria for the current study included age ≥ 40 years and no prior history of colon cancer.	N = 135	N = 135 G1: 45 G2: 45 G3: 45	N = 135 G1: 45 G2: 45 G3: 45	N = 135 G1: 45 G2: 45 G3: 45	
McCormack et al., 2011[4]	G1: control (no treatment control) G2: Prostate-Only (Net benefit) G3: Men's Health (Net benefit in context of other more beneficial services)	Convenience None Communities were selected using an in-depth comparison process involving sociodemographic and economic data	Inclusion: Men between 40 and 80 years of age and not diagnosed previously with prostate cancer	Overall N: 584 G1: 223 G2: 125 G3: 236	Overall N: 584 G1: 223 G2: 125 G3: 236	Overall N: 376 G1: 122 G2: 89 G3: 165	Overall N: 376 G1: 122 G2: 89 G3: 165	

Table F-2. Key question 3 sample characteristics, part 1 (continued)

Author, Year	Groups	Sampling Strategy, Unit of Randomization, Process of Randomization	Inclusion/ Exclusion Criteria	N Eligible	N Randomized	N Com-pleters	N Analyzed	Other Notes
Perneger et al., 2010[6] and 2011[5]	G1: control = minimal risk info, minimal benefit info G2: minimal risk info, moderate benefit info G3: minimal risk info, a lot of benefit info G4: moderate risk info, minimal benefit info G5: moderate risk info, moderate benefit info G6: moderate risk info, a lot of benefit info G7: a lot of risk info, minimal benefit info G8: a lot of risk info, moderate benefit info G9: a lot of risk info, a lot of benefit info Each participant received varying information about the benefits and harms of a screening test for an unnamed cancer.	NR Individual Computer generated random numbers	resident of canton Geneva and age 30-60	Overall N: 4670 G1-G9: NR	Overall N: 4670 G1-G9:NR	Overall N: 2333 G1: G2:	Overall N: 2327 G1-G9: NR	

Table F-2. Key question 3 sample characteristics, part 1 (continued)

Author, Year	Groups	Sampling Strategy, Unit of Randomization, Process of Randomization	Inclusion/ Exclusion Criteria	N Eligible	N Randomized	N Com-pleters	N Analyzed	Other Notes
Schwartz et al., 2011[7]	G1: control: No explanation of heart drug or heartburn drug G2: Nondirective explanation of heart drug or heartburn drug G3: Directive explanation of heart drug or heartburn drug Each participant sequentially randomized to 1 0f 3 groups for heart drug and then for heartburn drug	Sequential randomization Individual Central computerized random number generator	18+, member of the Knowledge Networks survey panel	Overall N: 4316 G1: G2:	Overall N: 2944 G1: surrogate outcome:981 new drug: 981 G2: surrogate outcome: 981 new drug: 982 G3: surrogate outcome: 982 new drug: 981	Overall N: 2944 G1: 981/981 G2: 981/982 G3: 982/981	Overall N: 2944 G1: 981/981 G2: 981/982 G3: 982/981	
Steginga et al., 2002[8]	G1: high risk/ low ambiguity G2: high risk/high ambiguity G3: low risk/low ambiguity G4: low risk/high ambiguity	All Individual NR	Inclusion: Men in urban and rural Queensland who were part of an ongoing study into men's health. Exclusion: Previously diagnosed with cancer (excluding non-melanoma skin cancer)	N = 919	Overall N: 705 G1: 173 G2: 175 G3: 178 G4: 179	N = 705 G1: 173 G2: 175 G3: 178 G4: 179	N = 705 G1: 173 G2: 175 G3: 178 G4: 179	

Abbreviations: ACCP = American College of Clinical Pharmacy; DX = diagnosis; G = group; GRADE = Grading of Recommendations Assessment, Development and Evaluation; N = number; NA = not applicable; NC = North Carolina; NICE = National Institute for Health and Clinical Excellence; NR = not reported

Table F-3. Key question 3 sample characteristics part 2

Author, Year	Groups	Mean Age	Female	Race, Ethnicity	Median Income	Insured	Edu-cation	Health literacy/ numeracy	Others Baseline Character-istics (^=signify-cant)	Other Notes
Akl et al., 2012[1]	G1: Strong and weak wording for or against guideline-supported behavior (ACCP): • "we recommend" • "we suggest" • "we suggest...not" "we recommend…not" G2: Strong and weak wording for or against guideline-supported behavior (NICE): • "clinicians should" • "clinicians might" • "clinicians might not" • "clinicians should not" G3: Strong and weak recommendations for or against guideline-supported behavior (GRADE): • "we recommend" • "we conditionally recommend" • "we conditionally recommend…not" • "we recommend…not"	M = 29.5	165, 48.1%	NR	NR	NR	NR	NR	NR	English native language (*possible different among groups) Country of med school graduation Graduated from medical school years ago Type of residency training Years of training

Table F-3. Key question 3 sample characteristics part 2 (continued)

Author, Year	Groups	Mean Age	Female	Race, Ethnicity	Median Income	Insured	Edu-cation	Health literacy/ numeracy	Others Baseline Character-istics (^=signify-cant)	Other Notes
Brewer et al., 2012[2]	G1: Percent + verbal descriptor (reference) G2: G1 + risk continuum graphic (reference) G3: G2 + confidence interval (precision) G4: G3 + risk score + graph (NA) G5: Oncotype DX report (precision) G5: Icon array (reference) Note: Each format presents the likelihood of recurrence of breast cancer	Range = 34-85	100%	White: 117, 89%	NR	<$60,000: 50, 41%	Insured: 123, 93%	College degree: 85, 64%	High health literacy (8 of 8 correct): 45, 79% [Only 57 patients received the health literacy assessment] High numeracy (3 of 3 correct): 42, 32%	Marriage status Worked for pay Received Onctotype DX test Received/ planning to receive treatment
Han et al., 2011[3] (Experiment 1)	G1: Point estimate in text G2: Point estimate as graph G3: Range in text G4: Range as graph NOTE: Each format tested before and after delivery of information about the populations average risk of colon cancer (6% versus 2 to 10%).	M = 52	"nearly equal male and female"	10% identified as non-white/Caucasian	NR	NR	NR	High school /GED or less = 74%	Number numeracy questions correct: 0 = 22.9% 1 = 33.8% 2 = 33.8% 3 = 9.6%	NR

Table F-3. Key question 3 sample characteristics part 2 (continued)

Author, Year	Groups	Mean Age	Female	Race, Ethnicity	Median Income	Insured	Edu-cation	Health literacy/ numeracy	Others Baseline Character-istics (^=signify-cant)	Other Notes
Han et al., 2011[3] (Experiment 2)	G1: Range in text (precision) G2: Range in text + solid bar graph (precision) G3: Range in text + blurred bar graph (precision)	M = 54	NR	8% identified as non-white/ Caucasian	NR	NR	NR	High school /GED or less = 26%	Number numeracy questions correct: 0 = 18.1% 1 = 34.7% 2 = 31.9% 3 = 15.3%	NR
McCormack et al., 2011[4]	G1: control (no treatment control) G2: Prostate-Only (Net benefit) G3: Men's Health (Net benefit in context of other more beneficial services)	63* significantly different, p<.05	0	36.3* significantly different, p<.05	NR	<$39,000- 34.0%* $40- 59,999- 21.0%* $60,000+- 34.3%* significantly different, p<.05	NR	High school or less: 12.8%* some college: 20.8%* college for more: 61.4%*	NR	PCP^ Recent PSA^ Knowledge of Prostate CA^ Cancer other than prostate Perceived risk of cancer Self-efficacy

Table F-3. Key question 3 sample characteristics part 2 (continued)

Author, Year	Groups	Mean Age	Female	Race, Ethnicity	Median Income	Insured	Edu-cation	Health literacy/ numeracy	Others Baseline Character-istics (^=signify-cant)	Other Notes
Perneger et al., 2010[6] and 2011[5]	G1: control = minimal risk info, minimal benefit info G2: minimal risk info, moderate benefit info G3: minimal risk info, a lot of benefit info G4: moderate risk info, minimal benefit info G5: moderate risk info, moderate benefit info G6: moderate risk info, a lot of benefit info G7: a lot of risk info, minimal benefit info G8: a lot of risk info, moderate benefit info G9: a lot of risk info, a lot of benefit info Each participant received varying information about the benefits and harms of a screening test for an unnamed cancer.	42.3 ±8.3^	1273(54.6)	NR	Swiss: 1655 (70.9) Other: 678 (29.1)	More than 4000 Swiss francs/month: 1722 (78.0)	NR	Higher: 1315 (57.1)	NR	Health status, MD visit in past 6 months, medical decision in past 6 months, Screening test in past 3 years^, Attitude toward screening^, Desire for info^, desire for autonomy, decision about hypothetical screening test^

Table F-3. Key question 3 sample characteristics part 2 (continued)

Author, Year	Groups	Mean Age	Female	Race, Ethnicity	Median Income	Insured	Edu-cation	Health literacy/ numeracy	Others Baseline Character-istics (^=signify-cant)	Other Notes
Schwartz et al., 2011[7]	G1: control: No explanation of heart drug or heartburn drug G2: Nondirective explanation of heart drug or heartburn drug G3: Directive explanation of heart drug or heartburn drug Each participant sequentially randomized to 1 0f 3 groups for heart drug and then for heartburn drug	46 (range, 18-93)	1531(52)*	White: 2041 (69)* Black, non-hispanic: 324 (11)* Hispanic: 383 (13)* Other: 191 (6.5)*	NR	NR	NR	<high school: 371*(12) high school grad: 930 (32)* some college: 824 (28)* college grad: 500 (17)* Postgrad degree: 324* (11)	NR	Region, Household income categories
Steginga et al., 2002[8]	G1: high risk/low ambiguity G2: high risk/high ambiguity G3: low risk/low ambiguity G4: low risk/high ambiguity	M = 54.19	0%	NR	NR	NR	NR	Primary school: 9% High school: 38% Vocational training: 18% Completed University studies: 35%	NR	Place of residence

Abbreviations: ACCP = American College of Clinical Pharmacy; CA = cancer; DX = diagnosis; GED = General Education Diploma; GRADE = Grading of Recommendations Assessment, Development and Evaluation; M = Mean; MD = medical doctor; N = number; NICE = National Institute for Health and Clinical Excellence; NR = not reported; PCP = primary care physician; PSA = prostate-specific antigen.

Table F-4. Key question 3 intervention descriptions

Author, Year	Groups	Comparators	Clinical Focus of the Evidence, Source of Information, Theory-Based, Prior Testing	Intervention Format, Delivery Agent, Intensity	Evidence Presentation	Message of Intervention	Other Notes
Akl et al., 2012[1]	G1: Strong and weak wording for or against guideline-supported behavior (ACCP): • "we recommend" • "we suggest" • "we suggest...not" "we recommend…not" G2: Strong and weak wording for or against guideline-supported behavior (NICE): • "clinicians should" • "clinicians might" • "clinicians might not" • "clinicians should not" G3: Strong and weak recommendations for or against guideline-supported behavior (GRADE): • "we recommend" • "we conditionally recommend" • "we conditionally recommend…not" • "we recommend…not"	Overall strength of evidence G3: Strong and weak recommendations for or against guideline-supported behavior: "we recommend" "we conditionally recommend" "we conditionally recommend…not" "we recommend…not" Overall strength of evidence G2: Strong and weak wording for or against guideline-supported behavior: "clinicians should" "clinicians might" "clinicians might not" "clinicians should not"" Overall strength of evidence G1: Strong and weak wording for or against guideline-supported behavior: "we recommend" "we suggest" "we suggest..not" "we recommend…not"	Two disease conditions: irritable bowel syndrome or congestive heart failure and a related hypothetical medication to treat those conditions. American College of Chest Physicians conferences on antithrombotic and thrombolytic therapy, the NICE, and the GRADE working group No No	Paper-based In-person 1 session with 2 scenarios, total time not reported	Qualitative	The participants read about a hypothetical drug and then read the statement that "a guidelines group that you trust has, based on recent randomized clinical trials, issued the following recommendation:" The participants read the recommendation (with different wording for the 3 groups) and then made a choice based on the recommendation.	

Table F-4. Key question 3 intervention descriptions (continued)

Author, Year	Groups	Comparators	Clinical Focus of the Evidence, Source of Information, Theory-Based, Prior Testing	Intervention Format, Delivery Agent, Intensity	Evidence Presentation	Message of Intervention	Other Notes
Brewer et al., 2012[2]	G1: Percent + verbal descriptor (reference) G2: G1 + risk continuum graphic (reference) G3: G2 + confidence interval (precision) G4: G3 + risk score + graph (NA) G5: Oncotype DX report (precision) G5: Icon array (reference) Note: Each format presents the likelihood of recurrence of breast cancer	Precision [Confidence Interval] G3: Described women's chance of breast cancer recurrence in 10 years as percentage (with a verbal interpretation of the percentage: low, medium, high) [Elements of G1] + a risk continuum graphic [G2] + text that reported a 95% CI with a verbal translation "chance of recurrence could be as low as 5% or as high as 9% for almost all 95% patients" Precision [Applicability + Precision] G5: Standard Onctotype DX report. Report included a statement that said "It is unknown whether the findings summarized in the clinical experience are applicable to patients with features different from those described" + recurrence score, a recurrence risk, a graph, a 95% CI, plain language risk categories,	Breast cancer recurrence Hhypothetical recurrence risk results derived from standard Oncotype DX reports developed by Genomic Health, Inc. Yes No	Paper-based In-person or by post 6 sessions, time not reported	G1: quantitative + qualitative G2, G3, G5: Quantitative + qualitative + graphical (combined) G6: graphical	Each vignette described a low, intermediate, or high chance of breast cancer recurrence in 10 years. Vignettes used 1 of 5 risk formats of various complexity that or a 6th format that used an icon array	

Table F-4. Key question 3 intervention descriptions (continued)

Author, Year	Groups	Comparators	Clinical Focus of the Evidence, Source of Information, Theory-Based, Prior Testing	Intervention Format, Delivery Agent, Intensity	Evidence Presentation	Message of Intervention	Other Notes
Brewer et al., 2012[2] (continued)		an assay description, and miscellaneous information about the test and Genomic health Usual Care G1: women's chance of breast cancer in 10 years described as a %. Interpreted with the evaluative labels: low, intermediate, or high chance. G2: G1 + risk continuum graphic (a horizontal bar chart from 0 to 100% partitioned into low, intermediate, high chance) G6: icon array depicting a woman's chances of breast cancer					

Table F-4. Key question 3 intervention descriptions (continued)

Author, Year	Groups	Comparators	Clinical Focus of the Evidence, Source of Information, Theory-Based, Prior Testing	Intervention Format, Delivery Agent, Intensity	Evidence Presentation	Message of Intervention	Other Notes
Han et al., 2011[3] (Experiment 1)	G1: Point estimate in text G2: Point estimate as graph G3: Range in text G4: Range as graph NOTE: Each format tested before and after delivery of information about the populations average risk of colon cancer (6% versus 2 to 10%).	Precision Point estimate of hypothetical risk of colon cancer Ambiguity ‘‘absent’’ was operationalized by communicating a hypothetical colorectal cancer risk estimate as a point estimate (‘‘Your chances of developing colon cancer in your lifetime are 9%’’). Precision Horizontal bar graph representing point estimate. Ambiguity ‘‘absent’’ + visual was operationalized by communicating a hypothetical colorectal cancer risk estimate as a point estimate (‘‘Your chances of developing colon cancer in your lifetime are 9%’’) along with a horizontal bar graph with solid border to depict the point estimate. Precision	Colorectal cancer risk NCI; evidence report Yes No	Web-based Online 1 session, time not reported	G1: quantitative G2: combined G3: quantitative G4: combined	Participants were told what their chances were of developing colon cancer in their lifetime. They were also provided with comparative risk information as a secondary pre-post test following the main assessment.	

Table F-4. Key question 3 intervention descriptions (continued)

Author, Year	Groups	Comparators	Clinical Focus of the Evidence, Source of Information, Theory-Based, Prior Testing	Intervention Format, Delivery Agent, Intensity	Evidence Presentation	Message of Intervention	Other Notes
Han et al., 2011[3] (Experiment 1) (continued)		G3: Range representing confidence intervals for hypothetical estimate of colon cancer. No point estimate provided. Ambiguity "present" was operationalized by communicating a hypothetical colorectal cancer risk estimate as a range (''Your chances of developing colon cancer in your lifetime are between 5% and 13%'') G4: Horizontal bar graph with solid borders depicting range. No point estimate provided. Ambiguity "present" + visual was the same as G3 except there was also a visual representation of the risk estimates (horizontal bar graphs with solid borders to depict point estimates and confidence intervals).					

Table F-4. Key question 3 intervention descriptions (continued)

Author, Year	Groups	Comparators	Clinical Focus of the Evidence, Source of Information, Theory-Based, Prior Testing	Intervention Format, Delivery Agent, Intensity	Evidence Presentation	Message of Intervention	Other Notes
Han et al., 2011[3] (Experiment 2)	G1: Range in text (precision) G2: Range in text + solid bar graph (precision) G3: Range in text + blurred bar graph (precision)	Precision Range in Text: Text only saying "Your chance of developing colon cancer in your lifetime are more likely between 5%-13%, but they could be higher or lower. Risk estimates are not exact." Precision Same text as G1 + solid bar graph range Precision Same text as G1 + bar graph range with blurred edges to reinforce the presence of ambiguity and the concept that probability distributions lack firm, categorical boundaries.	Colorectal cancer risk NCI; evidence report Yes Yes	Web-based Online 1 session, time not reported	G1: quantitative G2: combined G3: combined	The enhanced textual representation aimed at more explicitly emphasizing and describing the meaning of imprecision and uncertainty in risk predictions, whereas the enhanced visual representation depicted a confidence interval using a bar graph, but adding blurred edges to reinforce the presence of ambiguity and the concept that probability distributions lack firm, categorical boundaries.	

Table F-4. Key question 3 intervention descriptions (continued)

Author, Year	Groups	Comparators	Clinical Focus of the Evidence, Source of Information, Theory-Based, Prior Testing	Intervention Format, Delivery Agent, Intensity	Evidence Presentation	Message of Intervention	Other Notes
McCormack et al., 2011[4]	G1: control (no treatment control) G2: Prostate-Only (Net benefit) G3: Men's Health (Net benefit in context of other more beneficial services)	Net benefit Prostate Only: gave information on the uncertainty of PSA screening: There are fast and slow growing CA; PSA leads some men with slow-growing cancer to get treatment they don't need; about half of treated men will have side effects; men should decide if it's right for them Net Benefit [Net benefit + Framing] Men's Health (MH)- provides same info as G2 but ALSO gave additional info about other common preventive health screening for me and highlighted the certainty in benefit from these screening tests.	Prostate cancer screening USPSTF; guideline Yes Yes	In-person, video, web-based and paper Community physician #: 1 length: NA total time: 45 minutes	Combined	20 intervention large group sessions, with 10-30 male participants per session with an oral scripted presentation by a community physician followed by a question-and-answer session, a 20-minute video, a website, and print materials, including a tri-fold brochure, a 4-in x6-in. poster, and a shirt-pocket card decision aid	

Table F-4. Key question 3 intervention descriptions (continued)

Author, Year	Groups	Comparators	Clinical Focus of the Evidence, Source of Information, Theory-Based, Prior Testing	Intervention Format, Delivery Agent, Intensity	Evidence Presentation	Message of Intervention	Other Notes
Perneger et al., 2010[6] and 2011[5]	G1: control = minimal risk info, minimal benefit info G2: minimal risk info, moderate benefit info G3: minimal risk info, a lot of benefit info G4: moderate risk info, minimal benefit info G5: moderate risk info, moderate benefit info G6: moderate risk info, a lot of benefit info G7: a lot of risk info, minimal benefit info G8: a lot of risk info, moderate benefit info G9: a lot of risk info, a lot of benefit info Each participant received varying information about the benefits and harms of a screening test for an unnamed cancer.	Net benefit Hypothetical vignette with the amount of information about risks (3 levels) and benefits (3 levels). All versions included minimal descriptive information, and added information about benefits of the screening test (none, survival benefit, or survival benefit and reassurance) and associated risks (none, false positive results, false positive and false negative results) and were combined in a factorial design to yield 9 versions of the scenario More than minimal risk information (aggregated across G4-G9): • Moderate: false-positive results • A lot: false-positive and false-negative results	Unnamed cancer screening test Meta-analysis of efficacy of screening mammography in JAMA (although breast cancer not indicated as the condition in the vignette) Unclear Yes	Paper-based mailed survey of hypothetical vignette Postal #:1 length: NR total time: NR	Quantitative	Risk and benefits of hypothetical test	Based on prior work by same author

Table F-4. Key question 3 intervention descriptions (continued)

Author, Year	Groups	Comparators	Clinical Focus of the Evidence, Source of Information, Theory-Based, Prior Testing	Intervention Format, Delivery Agent, Intensity	Evidence Presentation	Message of Intervention	Other Notes
Perneger et al., 2010[6] and 2011[5] (continued)		More than minimal benefit information (aggregated across G2, G3, G5, G6, G8, G9): • Moderate: survival benefit • A lot: survival benefit and reassurance of testing of the screening test. Minimal risk information (aggregated across G1, G2, G3): none Minimal benefit information (aggregated across G1, G4, G7): none					
Schwartz et al., 2011[7]	G1: control: No explanation of heart drug or heartburn drug G2: Nondirective explanation of heart drug or heartburn drug G3: Directive explanation of heart drug or heartburn drug	Lack of directness G2: Nondirective explanation about the evidence: ("It takes time to establish the safety of new drugs."); received an explanation "surrogates do no always translate into patient outcomes" or "It takes time to establish the safety of new drugs" but NO DIRECTIVE advice	Medical prevention of heart attacks/lower cholesterol and treatment of heart burn NR No Yes	Web-based Web #: 1 length: NR total time: NR	Qualitative	Received either an explanation of "surrogates do no always translate into patient outcomes" or "It takes time to establish the safety of new drugs" and no directive advice or the explanation AND DIRECTIVE advice "Ask for a drug shown to reduce heart attacks" or "Ask for a drug with a longer track record"	Pre-tests for both logistics and qualitative changes-pg. 1464.

Table F-4. Key question 3 intervention descriptions (continued)

Author, Year	Groups	Comparators	Clinical Focus of the Evidence, Source of Information, Theory-Based, Prior Testing	Intervention Format, Delivery Agent, Intensity	Evidence Presentation	Message of Intervention	Other Notes
Schwartz et al., 2011[7] (continued)	Each participant sequentially randomized to 1 0f 3 groups for heart drug and then for heartburn drug	G2: Nondirective explanation about the evidence: ("It takes time to establish the safety of new drugs."); received an explanation "surrogates do no always translate into patient outcomes" or "It takes time to establish the safety of new drugs" but NO DIRECTIVE advice Lack of Directness G3: Nondirective explanation about the evidence and directive advice about what to do (“Ask for a drug with a longer track record"); received an explanation "surrogates do no always translate into patient outcomes" or "It takes time to establish the safety of new drugs" AND DIRECTIVE advice "Ask for a drug shown to reduce heart attacks" or "Ask for a drug with a longer track record"					

Table F-4. Key question 3 intervention descriptions (continued)

Author, Year	Groups	Comparators	Clinical Focus of the Evidence, Source of Information, Theory-Based, Prior Testing	Intervention Format, Delivery Agent, Intensity	Evidence Presentation	Message of Intervention	Other Notes
Schwartz et al., 2011[7] (continued)		Control G1: Control. No explanation about evidence.					
Steginga et al., 2002[8]	G1: high risk/ low ambiguity G2: high risk/high ambiguity G3: low risk/low ambiguity G4: low risk/high ambiguity	Precision High Ambiguity. G2 & G4 Ambiguity was manipulated by presenting an interval estimate of the likelihood of side effects. Risk was manipulated by changing the likelihood of prostate cancer cure from 80% (low risk) to 40% (high risk). Precision Low Ambiguity G1 & G3 Providing an unambiguous probability figure that was in the midrange of the interval estimate. Risk was manipulated by changing the likelihood of prostate cancer cure from 80% (low risk) to 40% (high risk).	Prostate cancer treatment for localized prostate cancer Not clear if it is previous research by Herr, 1997 and Shrader-Bogen et al., 1997, or if it is based on Australian Cancer Working Party Unclear Yes	Paper-based Postal 1 session, time not reported	Quantitative	A series of vignettes presenting information about localized prostate cancer, and its treatment including an operation or xray therapy. Side effects of each treatment are described. All side effects rates are presented using either high or low ambiguity. Treatment success is specified only in unambiguous terms (40% or 80%).	

Abbreviations: ACCP = American College of Clinical Pharmacy; CA = cancer; CI = confidence interval; DX = diagnosis; G = group; GRADE = Grading of Recommendations Assessment, Development and Evaluation; JAMA = Journal of the American Medical Association; NA = not applicable; NCI = National Cancer Institute; NICE = National Institute for Health and Clinical Excellence; NR = not reported; PSA = prostate-specific antigen; USPTF = US Preventive Services Task Force

Table F-5. Key question 3 first outcome

Author, Year	Groups	Outcome #1, Exact Measure Used	Timing of Measurement, Data Source	N Analyzed for this Outcome	Results by Group	Differences in Groups	Statistical Methods Used, Covariates Controlled for in Analysis
Akl et al., 2012[1]	G1: Strong and weak wording for or against guideline-supported behavior (ACCP): • "we recommend" • "we suggest" • "we suggest...not" "we recommend …not" G2: Strong and weak wording for or against guideline-supported behavior (NICE): • "clinicians should" • "clinicians might" • "clinicians might not" • "clinicians should not" G3: Strong and weak recommendations for or against guideline-supported behavior (GRADE): • "we recommend" • "we conditionally recommend" • "we conditionally recommend…not" • "we recommend…not"	Health-related decisions or behavior (applicable for general public/patients) Appropriate or inappropriate course of action based on choice of one of eight response options that correlate with the 4 strengths of recommendation and one of two decision-making styles (paternalism and shared decision making). To be considered appropriate, the chosen response option had to correspond to the strength and direction of recommendation presented in a table in the document (regardless of decision making style).	Immediate posttest Self-report	G1: 114 G2: 111 G3: 118	Appropriate Choices, by language category(%): "Strong for" G1: 7% G2: 9% G3: 7% "Weak for" G1: 77% G2: 46% G3: 61% "Weak against" G1: 32% G2: 55% G3: 64% "Strong against" G1: 49% G2: 42% G3: 51%	Difference: (p value) "Strong for" G1 vs. G2: -2%* G1 vs. G3: 0%* G2 vs. G3: 2%* p=0.91 "Weak for" G1 vs. G2: 31%* G1 vs. G3: 16%* G2 vs. G3: -15%* p=0.003 "Weak against" G1 vs. G2: -23%* G1 vs. G3: -32%* G2 vs. G3: -9%* p=0.002 "Strong against" G1 vs. G2: 7%* G1 vs. G3: -2%* G2 vs. G3: -9%* p =0.60	Chi-square and regression analysis (results of regression not reported other than to say that the analysis "confirmed the findings of the bivariate analyses regarding the association b/w wording and appropriate course of action.) For regression analysis (to determine what predicts appropriate actions): disease scenario, demographic characteristics, educational characteristics, and familiarity with the grading approaches Note: results in columns m and n are unadjusted.

Table F-5. Key question 3 first outcome (continued)

Author, Year	Groups	Outcome #1, Exact Measure Used	Timing of Measurement, Data Source	N Analyzed for this Outcome	Results by Group	Differences in Groups	Statistical Methods Used, Covariates Controlled for in Analysis
Brewer et al., 2012[2]	G1: Percent + verbal descriptor (reference) G2: G1 + risk continuum graphic (reference) G3: G2 + confidence interval (precision) G4: G3 + risk score + graph (NA) G5: Oncotype DX report (precision) G5: Icon array (reference) Note: Each format presents the likelihood of recurrence of breast cancer	Knowledge about the evidence Interpretation errors/accuracy of risk perception using 2 items. First item assessed whether women inaccurately identified the ‘‘gist’’ of the recurrence risk presented Response options were ‘‘low chance,’’ ‘‘intermediate chance,’’ or ‘‘high chance.’’ The second item assessed whether women inaccurately identified the verbatim recurrence risk. Response options were 0–100%.	Immediate post-test after each of the 6 vignettes Self-report	Gist error G1: 133 G2: 133 G3: 133 G4: 114 G4: 133 G6: 133 Verbatim errors G1: 130 G2: 133 G3: 129 G4: 111 G5: 128 G6: 127	Gist errors G1: 13% G2: 6% G3: 5% G5: 17% G6: 16%	Gist errors G1 vs. G3: +8%*, p=NR G2 vs. G3: +1%, p=NR G6 vs. G3: +11%, p=NR G1 vs. G5: 4%*; OR:0.57(0.31 to 1.06), p=ns G2 vs. G5: 11%*; OR: 0.27 (0.12 to 0.58), p<0.001 G3 vs. G5: 12%* OR:0.23(0.10 to 0.52), p<0.001 G5 vs. G6: 1%*; OR:0.79(0.44 to 1.44),p=ns	Generalized estimating equations Risk serial position

Table F-5. Key question 3 first outcome (continued)

Author, Year	Groups	Outcome #1, Exact Measure Used	Timing of Measurement, Data Source	N Analyzed for this Outcome	Results by Group	Differences in Groups	Statistical Methods Used, Covariates Controlled for in Analysis
Han et al., 2011[3] (Experiment 1)	G1: Point estimate in text G2: Point estimate as graph G3: Range in text G4: Range as graph NOTE: Each format tested before and after delivery of information about the populations average risk of colon cancer (6% versus 2 to 10%).	Knowledge about the evidence Perceived risk of getting colon cancer in your lifetime. Measured using 2 items that were averaged: 1) Based on these results, how would you describe your chances of developing colon cancer in your lifetime? very low to very high on 5-point likert; 2) If I received these results, I would feel that I'm going to get colon cancer. strongly disagree to strongly agree on 5 point likert.	Immediate post-test Self-report	N = 240 G1: 60 G2: 60 G3: 60 G4: 60	Before comparative risk information (Not reported in text, values estimated from Figure 3) Text G1: 1.7* Graph G2: 2.1* Text G3: 2.0* Graph G4: 1.6* *estimated from Figure After comparative risk information (perceived risk change, ranging from 0-1, estimated from Figure 5) Text G1: 0.70* Graph G2: 0.15* Text G3: 0.12* Graph G4: 0.35* *estimated from Figure	Before comparative risk information No significant main effect of ambiguity or representational format on perceived risk. G3-G1: +0.3, p=NS G4-G2: -0.5, p=NS Ambiguity × Representational Format was significant (F (1, 231) = 9.08, p=0.003) After comparative risk information Significant main effect of ambiguity and representational format on perceived risk (F (1, 229) = 4.86, p=0.03). Significant 3-way interaction of ambiguity x format x comparative risk information (Wilks's l=:92, Fð2, 227) = 9.41, p<0.001). G3-G1: -0.58, p=0.03 G4-G2: +0.20, p=0.03	MANOVA and ANOVA Dispositional optimism and numeracy

Table F-5. Key question 3 first outcome (continued)

Author, Year	Groups	Outcome #1, Exact Measure Used	Timing of Measurement, Data Source	N Analyzed for this Outcome	Results by Group	Differences in Groups	Statistical Methods Used, Covariates Controlled for in Analysis
Han et al., 2011[3] (Experiment 2)	G1: Range in text (precision) G2: Range in text + solid bar graph (precision) G3: Range in text + blurred bar graph (precision)	Knowledge about the evidence Perceived risk of getting colon cancer in your lifetime. Measured using 2 items that were averaged: 1) Based on these results, how would you describe your chances of developing colon cancer in your lifetime? very low to very high on 5-point likert; 2) If I received these results, I would feel that I'm going to get colon cancer. Strongly disagree to strongly agree on 5 point likert.	Immediate post-test Self-report	N = 135 G1: 45 G2: 45 G3: 45	NR	Effect size NR, NS.	MANOVA and ANOVA Dispositional optimism and numeracy

Table F-5. Key question 3 first outcome (continued)

Author, Year	Groups	Outcome #1, Exact Measure Used	Timing of Measurement, Data Source	N Analyzed for this Outcome	Results by Group	Differences in Groups	Statistical Methods Used, Covariates Controlled for in Analysis
McCormack et al., 2011[4]	G1: control (no treatment control) G2: Prostate-Only (Net benefit) G3: Men's Health (Net benefit in context of other more beneficial services)	Knowledge about the evidence 10 "demonstrated-knowledge" questions were used to calculate participant's knowledge of the contents of the interventions. A knowledge index score was computed and ranged from 0 to 10. (10 = all correct)	Baseline, 6 months, 12 months Self-report	Overall N: 376 G1: 122 G2: 89 G3: 165	Mean knowledge scores at 6 months (range 0-10): G1: 3.6 G2: 5.1 G3: 4.9 Mean knowledge scores at 12 months (range 0-10): G1: 3.7 G2: 4.5 G3: 4.5 Active Involvement in Decision Making, %: G1: 75% G2: 79% G3: 78% PSA screening at 12 month follow-up, %: G1: 64% G2: 71% G3: 61%	Mean knowledge scores at 6 months, absolute difference: G3-G1: 1.3*, p=NR G2-G1:1.5*, p=NR G3-G2: 0.2*, p=NR Mean knowledge score increase at 12 months from baseline, absolute difference: G3-G1: +1.5* , p<0.001 G2-G1: +0.9* , p<0.05 Absolute difference in decision making involvement: G3-G1: + 3%, p=0.045 G2-G1: + 4%, p=0.064 G3-G2: -1%, p=NR Absolute difference in PSA screening (12 months) G3-G1: -3%*, p=NR G2-G1: 7%*, p=NR G3-G2: -10%*, p=NR	GEE modeling Education, marital status, prior PSA testing, health status, health literacy, race Higher education, being married, ever had a PSA test, excellent/very good self-reported health (versus fair/poor), and greater health literacy were associated with higher knowledge scores. Lower knowledge scores were associated with being Black versus White.

Table F-5. Key question 3 first outcome (continued)

Author, Year	Groups	Outcome #1, Exact Measure Used	Timing of Measurement, Data Source	N Analyzed for this Outcome	Results by Group	Differences in Groups	Statistical Methods Used, Covariates Controlled for in Analysis
Perneger et., al., 2010[6] and 2011[5]	G1: control = minimal risk info, minimal benefit info G2: minimal risk info, moderate benefit info G3: minimal risk info, a lot of benefit info G4: moderate risk info, minimal benefit info G5: moderate risk info, moderate benefit info G6: moderate risk info, a lot of benefit info G7: a lot of risk info, minimal benefit info G8: a lot of risk info, moderate benefit info G9: a lot of risk info, a lot of benefit info Each participant received varying information about the benefits and harms of a screening test for an unnamed cancer.	Health-related decisions or behavior (applicable for general public/patients) Decision Evaluation: Combined score ranging from 0 (lowest) to 100 (highest) based on 3 items from a modified decisional conflict scale and 3 items from the satisfaction with decision scale	Given in postal survey Self-report	N=2333	DECISION SATISFACTION: G1:85.9 (17.4) G2: 86.4 (14.4) G3: 86.1 (16.0) G4: 79.2 (19.5) G5: 79.4 (18.1) G6: 81.8 (17.0) G7: 79.3 (17.6) G8: 81.1 (18.8) G9: 83.2 (17.2) Mean decision satisfaction: Minimal risk, aggregate benefit: 85.9 Mod risk, aggregate benefit: 80.4 Lot of risk, aggregate benefit: 81.2 Minimal benefit, aggregate risk: 81,4 Moderate benefit, aggregate risk: 82.5 Lot of benefit, aggregate risk: 83.6: % Test Refusal: Minimal risk, aggregate benefit: 8.8 Minimal benefit, aggregate risk: 16.6	Adjusted absolute difference in Decision Satisfaction: More than minimal vs. minimal risk: -5.1 (-6.6, -3.6) More than minimal vs. minimal benefit: 1.1 (-0.4 to 3.6) OR for test refusal (compared to minimal information): Minimal risk info: 1.0 Moderate risk info (FP): 2.5 (1.8 to 3.4) Lot of risk info (FP + FN): 3.0 (2.2 to 4.2) Minimal benefit info: 1.0 Mod benefit info (survival): 1.0 (0.7 to 1.3) Level of benefit info (survival and reassurance): 1.0 (0.7 to 1.3)	Two-way ANOVA Age, screening in past 3 years, desire for information

Table F-5. Key question 3 first outcome (continued)

Author, Year	Groups	Outcome #1, Exact Measure Used	Timing of Measurement, Data Source	N Analyzed for this Outcome	Results by Group	Differences in Groups	Statistical Methods Used, Covariates Controlled for in Analysis
Schwartz et al., 2011[7]	G1: control: No explanation of heart drug or heartburn drug G2: Nondirective explanation of heart drug or heartburn drug G3: Directive explanation of heart drug or heartburn drug Each participant sequentially randomized to 1 0f 3 groups for heart drug and then for heartburn drug	Health-related decisions or behavior (applicable for general public/patients) Choice of the better drug (more distal outcomes)	Immediately following intervention Self-report	Overall N: 2944 G1: 981 G2: 981 G3: 982	G1: 59% G2: 71% G3: 71%	G1-G2 Difference: 12 % 95% CI: 7-18 p=NR G1-G3 Difference: 12 % 95% CI: 7-18 p=NR	Unclear- use the SVY series of commands- and post-estimation commands for CI None

Table F-5. Key question 3 first outcome (continued)

Author, Year	Groups	Outcome #1, Exact Measure Used	Timing of Measurement, Data Source	N Analyzed for this Outcome	Results by Group	Differences in Groups	Statistical Methods Used, Covariates Controlled for in Analysis
Steginga et al., 2002[8]	G1: high risk/ low ambiguity G2: high risk/high ambiguity G3: low risk/low ambiguity G4: low risk/high ambiguity	Health-related decisions or behavior (applicable for general public/patients) Preferences for involvement in decision making measured using Desire for Involvement in Decision Making scale, which asked participants how they would prefer decisions about their treatment to be made if they were diagnosed with prostate cancer on a Likert-type scale of 1 (The doctor alone) to 5 (You alone)	Immediate posttest Self-report	N = 705 G1: 173 G2: 175 G3: 178 G4: 179	M (SD) G1: 3.23 (.76) G2: 3.31 (.73) G3: 3.25 (.73) G4: 3.23 (.82)	No main or interaction effects of ambiguity and risk on preference for decisional style, p's >0.35	ANOVA NR

Abbreviations: ACCP = American College of Clinical Pharmacy; ANOVA = ANalysis Of Variance; b/t = between; DX = diagnosis; FN = false negative; FN = false positive; G = group; GEE = generalized estimating equations method; GRADE = Grading of Recommendations Assessment, Development and Evaluation; M = Mean; MANOVA = Multivariate analysis of variance; NA = not applicable; NICE = National Institute for Health and Clinical Excellence; NR = not reported; NS = not significant; OR = odds ratio; PSA = prostate-specific antigen; SD = standard deviation; SVY = survey; vs. = versus

Table F-6. Key question 3 second outcome

Author, Year	Groups	Outcome #2, Exact Measure Used	Timing of Measurement, Data Source	N analyzed for This Outcome	Results by Group	Differences in Groups	Statistical Methods Used, Covariates Controlled for in Analysis
McCormack et al., 2011[4]	G1: control (no treatment control) G2: Prostate-Only (Net benefit) G3: Men's Health (Net benefit in context of other more beneficial services)	Health-related decisions or behavior (applicable for general public/patients) PSA in the 12 months since the intervention. Participants were asked if they had a PSA test in the last year	12 months Self-report	Overall N: 355 G1: 118* G2: 85* G3: 152*	G1: 76, 64% G2: 60, 71% G3: 93, 61%	Difference G3-G2: -10%* P=NR G2-G1: 7%* P=NR G3-G1: 3%* P=NR	Absolute differences & Logistic regression (not reported) NR

Table F-6. Key question 3 second outcome (continued)

Author, Year	Groups	Outcome #2, Exact Measure Used	Timing of Measurement, Data Source	N analyzed for This Outcome	Results by Group	Differences in Groups	Statistical Methods Used, Covariates Controlled for in Analysis
Perneger et., al., 2010[6] and 2011[5]	G1: control = minimal risk info, minimal benefit info G2: minimal risk info, moderate benefit info G3: minimal risk info, a lot of benefit info G4: moderate risk info, minimal benefit info G5: moderate risk info, moderate benefit info G6: moderate risk info, a lot of benefit info G7: a lot of risk info, minimal benefit info G8: a lot of risk info, moderate benefit info G9: a lot of risk info, a lot of benefit info Each participant received varying information about the benefits and harms of a screening test for an unnamed cancer.	Health-related decisions or behavior (applicable for general public/patients) Test refusals (yes/no): %	Given in postal survey Self-report	N=2333	TEST REFUSALS: G1: 11.6% G2: 5.3% G3: 10% G4: 16.1% G5: 21.7% G6: 18.8% G7: 22.3% G8: 23.2% G9: 20%	OR for test refusal Minimal risk info: 1.0 Moderate risk info (FP): 2.5 (1.8 to 3.4) Lot of risk info (FP + FN): 3.0 (2.2 to 4.2) Minimal benefit info: 1.0 Mod benefit info (survival): 1.0 (0.7 to 1.3) Level of benefit info (survival and reassurance): 1.0 (0.7 to 1.3)	Chi square tests, logistic regression Risk benefit information, health status, medical decision in past 6 month, screening in past 3 years, attitude toward screening, desire for information, desire for autonomy

Table F-6. Key question 3 second outcome (continued)

Author, Year	Groups	Outcome #2, Exact Measure Used	Timing of Measurement, Data Source	N analyzed for This Outcome	Results by Group	Differences in Groups	Statistical Methods Used, Covariates Controlled for in Analysis
Schwartz et al., 2011[7]	G1: control: No explanation of heart drug or heartburn drug G2: Nondirective explanation of heart drug or heartburn drug G3: Directive explanation of heart drug or heartburn drug Each participant sequentially randomized to 1 0f 3 groups for heart drug and then for heartburn drug	Health-related decisions or behavior (applicable for general public/patients) Choice of the better drug- one that had been on the market longer	Immediately following intervention Self-report	Overall N: 2944 G1: 981 G2: 982 G3: 981	G1: 34% G2: 53% G3: 53%	G1-G2 Difference: 19 % 95% CI: 13-124 p=NR G1-G3 Difference: 19 % 95% CI: 13-24 P= NR	Unclear- use the SVY series of commands- and post-estimation commands for CI Used post-stratification weights to account for sampling strategy which adjusted for demographic characteristics

Abbreviations: FN = false negative; FN = false positive; G = group; N = number; OR = odds ratio; PSA = prostate-specific antigen; SVY = survey.