Technical Review
Number 16

Updating Systematic Reviews

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. 290-02-0021

Prepared by:
University of Ottawa Evidence-based Practice Center, Ottawa, Canada

Investigators
Kaveh G. Shojania, M.D.
Margaret Sampson, M.L.I.S.
Mohammed T. Ansari, M.B.B.S., M.Med.Sc., M.Phil.
Jun Ji, MD, M.H.A.
Chantelle Garritty, B.A., D.C.S., M.Sc.(c)
Steve Doucette, M.Sc.
Tamara Rader, M.L.I.S.
David Moher, Ph.D.

AHRQ Publication No. 07-0087
September 2007

This report is based on research conducted by the University of Ottawa Evidence-based Practice Center (EPC) under contract to the Agency for Healthcare Research and Quality (AHRQ), Rockville, MD (Contract No. 290-02-0021). The findings and conclusions in this document are those of the author(s), who are responsible for its content, and do not necessarily represent the views of AHRQ. 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.

Suggested Citation:
Shojania KG, Sampson M, Ansari MT, Ji J, Garritty C, Rader T, Moher D. Updating Systematic Reviews. Technical Review No. 16. (Prepared by the University of Ottawa Evidence-based Practice Center under Contract No. 290-02-0017.) AHRQ Publication No. 07-0087. Rockville, MD: Agency for Healthcare Research and Quality. September 2007.

No investigators have any affiliations or financial involvement (e.g., employment, consultancies, honoraria, stock options, expert testimony, grants or patents received or pending, or royalties) that conflict with material presented in this report.

Preface

The Agency for Healthcare Research and Quality (AHRQ), through its Evidence-Based Practice Centers (EPCs), sponsors the development of evidence reports and technology assessments to assist public- and private-sector organizations in their efforts to improve the quality of health care in the United States. The reports and assessments provide organizations with comprehensive, science-based information on common, costly medical conditions and new health care technologies. The EPCs systematically review the relevant scientific literature on topics assigned to them by AHRQ and conduct additional analyses when appropriate prior to developing their reports and assessments.

To bring the broadest range of experts into the development of evidence reports and health technology assessments, AHRQ encourages the EPCs to form partnerships and enter into collaborations with other medical and research organizations. The EPCs work with these partner organizations to ensure that the evidence reports and technology assessments they produce will become building blocks for health care quality improvement projects throughout the Nation. The reports undergo peer review prior to their release.

AHRQ expects that the EPC evidence reports and technology assessments will inform individual health plans, providers, and purchasers as well as the health care system as a whole by providing important information to help improve health care quality.

We welcome comments on this evidence report. 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.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

Beth A. Collins Sharp, Ph.D., R.N.
Director, EPC Program
Agency for Healthcare Research and Quality

David Atkins, M.D., M.P.H.
Chief Medical Officer
Agency for Healthcare Research and Quality

Acknowledgments

We would like to acknowledge the guidance and expertise contributed to this project by the Technical Expert Panel members:

- David Atkins – Chief Medical Officer, Agency for Healthcare Research and Quality
- Paul Shekelle – RAND (USA)
- Evelyn P. Whitlock – Kaiser Permanente Center for Health Research (USA)
- Cynthia Mulrow – University of Texas, San Antonio and Annals of Internal Medicine (USA)
- Doug Altman – Center for Statistics in Medicine (UK)
- Martin Eccles – Center for Health Services Research, Newcastle University (UK)
- P.J. Devereaux – McMaster University Health Sciences Centre (CAN)

We would like to acknowledge the contributions of members of the University of Ottawa Evidence-based Practice Center and others. Many of these researchers made significant contributions to portions of the project.

- Lorri Puil, Director of Systematic Reviews, Chalmers Research Group, for support and supervision of research staff
- Keith O'Rourke and Nicholas J Barrowman, for statistical advice
- Andrea Tricco for guidance with survey methods
- Alexander Tsertsvadze, for early methodological discussions and assistance with screening
- Tanya Armour, for mentorship of team members
- Jessie McGowan, for guidance with search method development
- Raymond Daniel, for search assistance and document acquisition
- Alla E. Iansavichene, for search assistance and data quality management
- Mary Ocampo, for technical and administrative assistance
- Alison Jenkins, for development of the meta-analytic Excel worksheet

Structured Abstract

Background: Systematic reviews are often advocated as the best source of evidence to guide both clinical decisions and healthcare policy, yet we know very little about the extent to which they require updating.

Objectives:

- To estimate the average time to changes in evidence sufficiently important to warrant updating systematic reviews (referred to as the survival time) and to identify any characteristics that increase or decrease these survival times.
- To determine the performance characteristics of various surveillance protocols to identify important new evidence.
- To assess the utility of rates and patterns of growth for evidence within clinical areas as predictors of updating needs.
- To establish typical timeframes for the production and publication of systematic reviews in order to assess the extent to which they impact survival time (e.g., whether or not delays in the peer review and publication processes substantially shorten the time in the public domain before new evidence requires updating of a given systematic review).
- To characterize current updating practices and policies of agencies that sponsor systematic reviews.

Design: Survival analysis for a cohort of 100 quantitative systematic reviews that were indexed in *ACP Journal Club* with an accompanying commentary; supplementary sample of Cochrane reviews meeting the same criteria and AHRQ evidence reports; internet-based survey of agencies that sponsor or undertake systematic reviews.

Sample: Eligible reviews evaluated the clinical benefit or harm of a specific (class of) drug, device, or procedure, were originally published between 1995 and 2005, and included at least one quantitative synthesis result in the form of an odds ratio, relative risk, risk difference, or mean difference. For the survey of updating policies and practices, we contacted 22 organizations that are well-known to produce or fund systematic reviews (including 12 AHRQ Evidence-based Practice Centers).

Data sources: Systematic reviews indexed in ACP J Club and eligible new trials identified through five search protocols.

Measurements: Quantitative signals for updating consisted of changes in statistical significance or a relative change in effect magnitude of at least 50 percent involving one of the primary outcomes of the original systematic review or any mortality outcome. These signals were assessed by comparing the original meta-analytic results with updated results that included eligible new trials. Qualitative signals included substantial differences in characterizations of effectiveness, new information about harm, emergence of superior alternative treatments, and important caveats about the previously reported findings that would affect clinical decisionmaking.

The primary outcome of interest was the occurrence of either a qualitative or quantitative signal for updating the original systematic review. We also assessed the occurrence of a signal for updating within 2 years of publication, as some sources (e.g., The Cochrane Library) currently recommend updating systematic reviews every two years.

The survey measured existing updating policies, current strategies in use, and additional perceptions related to the updating process from the 18 organizations that responded.

Results: The cohort of 100 systematic reviews included a median of 13 studies (inter-quartile range: 8 to 21) and 2663 participants (inter-quartile range: 1281 to 8371) per review. A qualitative or quantitative signal for updating occurred for 57 systematic reviews. Median survival free of a signal for updating was 5.5 years (95% confidence interval [CI]: 4.6-7.6), but in 23 cases (95% CI: 15% to 33%), a signal for updating occurred in less than 2 years, and in 15 cases (95% CI: 9% to 24%) the signal occurred in less than 1 year. In 7 cases (95% CI: 3% to 14%), a signal had already occurred at the time of publication of the original review. Shorter survival was associated with cardiovascular medicine (hazard ratio of 3.26, 95% CI: 1.71 to 6.21; p =0.0003), heterogeneity in the original review (hazard ratio of 2.23, 95% CI: 1.22 to 4.09; p =0.01), and having a new trial larger than the previous largest trial (hazard ratio of 1.08, 95% CI: 1.02 to 1.15; p =0.01). Systematic reviews with more than the median of 13 included studies had increased survival (hazard ratio of 0.55; 95% CI: 0.31 to 0.98; p =0.04). No feature of the original review significantly predicted a signal for updating occurring within 2 years of publication.

Median time from the final search date to indexing 1.4 years (inter-quartile range; 0.96-2.0 years). Lags from search to publication were shortest for Cochrane reviews (median 0.6 years, inter-quartile range: 0.42-1.25) and longest for journal reviews (median 1.3 years; inter-quartile range: 0.84-1.77), with technical reports falling in between (median 1.1 years; inter-quartile range: 0.87-1.42) (Kruskal Wallis $\chi 2$ 11.24, p =0.004).

Of the five search protocols tested for their effectiveness in identifying eligible new trials, the combination with the highest recall and lowest screening burden were the strategy that used the PubMed Related Articles feature (applied to the three newest and three largest trials included in the original review) and the strategy involved submitting a subject search (based on population and intervention) to the Clinical Query filter for therapy. This combination identified most new signaling evidence with median screening burden of 71 new records per review.

For the survey of organizations involved in producing or funding systematic reviews, we received responses from 19 (86%) of the 22 organizations contacted. Approximately two thirds (68%) of respondents identified themselves as producers of systematic reviews and an additional 21% identified themselves as both funders and producers of systematic reviews. Only two respondents (11%) characterized themselves solely as funders of systematic reviews.

Approximately 80% of respondents characterized the importance of updating as ‘high’ or ‘very high’, although 68% acknowledged not having any formal policies for updating in place. Approximately two thirds (13/19; 68%) of respondents reported that over 20% of the reviews they commission or produce are out of date, and 32% respondents (6/19) reported that at least 50% of their reviews were out of date. Barriers to updating identified by respondents included lack of appropriate methodologies, resource constraints, lack of academic credit, and limited publishing formats. The majority of the sample (16/19; 84%) indicated they ‘somewhat’ to

'strongly' favor the development of a central registry, analogous to efforts within the clinical trials community, to coordinate updating activities across agencies and review groups.

Conclusions: In a cohort of high quality systematic reviews directly relevant to clinical practice, signals for updating occurred frequently and within relatively short timelines. A number of features significantly affected survival, but none significantly predicted the need for updating within 2 years.

Currently, definitive methods about the frequency of updating cannot be made. Blanket recommendation such as every two years will miss a substantial number of important signals for updating that occur within shorter time lines, but more frequent updates will expend substantial resources. Methods for identifying reviews in need of updating based on surveillance for new evidence hold more promise than relying on features of the original review to identify reviews likely to need updating within a short time, but such approaches will require further investigation. Several of the methods tested were feasible, yielding good recall of relevant new evidence with modest screening burdens.

The majority of organizations engaged in the funding or production of systematic reviews view the importance of updating systematic reviews as high to very high. Despite this recognition, most organizations report having no formal policy in place for updating previous systematic reviews. Slightly less than half of organizations performed periodic literature searches to identify new evidence, but searching frequencies varied widely, from monthly to every two years.

If systematic reviews are to achieve their stated goal of providing the best evidence to inform clinical decision making and healthcare policy, issues related to identifying reviews in need of updating will require much greater attention. In the meantime, publishers of systematic reviews should consider a policy of requiring authors to update searches performed over 12 months prior to submission. And, users of systematic reviews need to recognize that important new evidence can appear within short timelines. When considering the results of a particular systematic review, users should search for more recent reviews or trials to see if any exist and determine if the results are consistent with the previous review.

Contents

Figures

Tables

Appendixes

Appendixes and Evidence Tables are provided electronically at http://www.ahrq.gov/downloads/pub/evidence/pdf/sysrev/sysrev.pdf

Technical Review

Chapter 1. Introduction

Overview

Systematic reviews are being published with increasing frequency. One recent estimate is that 2500 new systematic reviews are published annually.[1] This high volume reflects the several key roles played by systematic reviews: synthesizing for clinicians the evidence addressing a given topic; providing the foundation for the development of clinical practice guidelines; and informing cost-effectiveness analyses and policy decisions. In addition, some granting agencies now require that researchers include systematic reviews in grant applications to support the rationale for proposed new research. Fulfilling these important roles requires that systematic reviews be up to date. However, almost no empiric data indicate the extent to which systematic reviews require updating or the intervals at which updates should be performed. Given the paucity of data on this important topic, and with funding through the Evidence-based Practice Center, we set out to generate data to help inform optimal approaches to updating.

Background

The annual publication of systematic reviews has increased dramatically in recent years,[2,2] with an estimated 2500 new systematic reviews published per year.[1] The Cochrane Collaboration sets a goal of updating its reviews every 2 years, with the result that 38% of new Cochrane reviews represent updates of previous reviews (typically conducted by the same authors).[1] By contrast, only 2% of systematic reviews published in all other journals represent updates of previous reviews (whether conducted by the same authors or not). [1] Currently we have no way of knowing if either of these numbers adequately matches the true need for maintaining the currency of published systematic reviews.

Updating a previous systematic review requires resources. Even when the same authors update the review, carrying out the search and screening processes for potentially eligible new articles, abstracting new articles that are identified, and analyzing the results all take time. Agencies that commission and host systematic reviews thus face the challenge of how best to allocate resources between funding new reviews and supporting the maintenance of existing reviews. Even the process of determining if an aging review still represents a valid, even if no longer completely up to date, synthesis that can safely be left in the public domain, or new evidence alters the findings of the review to such an extent that withdrawing or archiving the review represents the preferred choice.

The goals of this effort are to determine the extent to which systematic reviews require updating and to ascertain approaches towards updating by organizations engaged in the funding or production of systematic reviews. This work builds upon and complements several pieces of work in which members of this team have participated over the past 24 months, including development of the first formal definition of updating to appear in the literature,[3] and a systematic review of existing methodologies for updating.[4] The evidence base proved to be quite limited, with a small body of literature on cumulative meta-analysis and the volume of new evidence needed to overturn previous meta-analytic results,[5,6] one evaluation of the shelf-life of

17 clinical practice guidelines produced for the Agency for Healthcare Research and Quality,[7] and some suggested approaches to updating from the Cochrane Collaboration.[8-10]

Given this paucity of literature informing approaches to updating, we set out to evaluate how soon systematic reviews require updating, and how best to detect the need for updating through empirical study of a cohort of quantitative systematic reviews. We also sought to determine current practices with respect to updating through consultation with agencies involved with funding or producing systematic reviews.

Key Questions Addressed in This Report

This evidence report aims to provide empiric data that address the following questions about updating systematic reviews.

1. How quickly do systematic reviews become out of date (i.e., what is the average 'shelf life' or ‘survival time’ of systematic reviews)?
2. Do any features of a given systematic review (including characteristics related to the content area, features of the included studies, and the nature of the results) increase or decrease ‘survival’ time?
3. What impact do publication time lags have on survival times for systematic reviews and what strategies can maximize the currency of systematic reviews at the time of publication?
4. Is the pattern of growth of evidence (e.g., the velocity of trial publication) within clinical areas predictive of the need to update?
5. Of agencies or organizations that fund or conduct systematic reviews, what are their existing policies or practices regarding updating systematic reviews?

Questions 1 to 4 are the main focus of this evidence report. They are addressed through identification and evaluation of new evidence pertaining to each of 100 systematic reviews, with an augmented sample of Cochrane reviews and AHRQ evidence reports for Question 3. Methods and main finding are described here in this extended overview. Each question is fully reported in a journal manuscript. The fifth question was addressed through a survey of 9 agencies as well as the EPCs, who also responded to several additional questions of particular interest to the AHRQ’s EPC Program. The survey served as a pilot for a larger survey not conducted under the auspices of AHRQ. The material from the pilot will be reported in a separate manuscript. Methods and major results of the pilot survey are presented here; the larger survey will be reported in a subsequent manuscript.

Chapter 2. Methods

Study Identification

Search Strategy

The first four questions are explored through a cohort of 100 systematic reviews identified through a search of *ACP Journal Club* database on Ovid, undertaken January 31, 2006. The search to identify candidates screened for inclusion in the cohort was:

1. review$.ti.
2. meta-analy$.mp.
3. data sources.ab.
4. (search$ or MEDLINE®).ab.
5. or/1-4
6. limit 5 to articles with commentary

Additional Cochrane reviews included for Question 3 were identified through the same search. Additional AHRQ reports used for Question 3 were identified through PubMed® with the query "Evid Rep Technol Assess (Summ)"[Journal:__jrid21544]. Searches were undertaken April 10, 2006.

Eligibility Criteria

The time to important changes in evidence might vary depending on a number of factors, including the type of question posed by the original review (e.g., therapeutic, diagnostic, prognostic, or health policy), the type of studies included (e.g., randomized controlled trials, observational studies), and whether or not the systematic review provided quantitative synthesis. In the interest of reducing potential sources of variation, we focused on systematic reviews that evaluated the clinical benefit or harm of a specific (class of) drug, device, or procedure and provided quantitative synthesis that included a point estimate and 95% confidence interval for at least one clinical outcome (disease endpoint, functional status, mortality) or established intermediate outcome (e.g., blood pressure, glycemic control, standard instrument for measuring disease activity, such as a depression scale). We excluded evaluations of alternative and complementary medicines, as well as educational and behavioral interventions.

Further eligibility requirements were as follows:

- Publication from 1995 to 2005 (but with search date no later than Dec 31, 2004 to ensure at least one full year for new evidence to appear)
- Reporting of at least one conventional meta-analytic estimate of treatment benefit or harm. We excluded individual patient data meta-analyses, meta-regressions, and indirect meta-analyses because of the difficulty of determining whether or not data from new trials would alter previous quantitative results.

- Included at least one randomized controlled trial; other eligible designs were restricted to quasi-randomized or controlled clinical trials (CCTs).
- Meta-analytic outcomes reported in the form of a relative risk, odds ratio, or absolute risk difference for binary outcomes and weighted mean differences for continuous outcomes. We excluded standardized effect sizes to avoid the complexity of assessing candidate new data reported using different outcome scales to determine if they would have met the authors' criteria for incorporation into the standardized effect measure in the original review.

We used as our sampling frame systematic reviews that were selected for commentaries in *ACP Journal Club*, a bimonthly publication of the American College of Physicians that aims "to select from the biomedical literature articles that report original studies and systematic reviews that warrant immediate attention by physicians attempting to keep pace with important advances in internal medicine." [11] The article selection process involves "reliable application of explicit criteria for scientific merit, followed by assessment of relevance to medical practice by clinical specialists." Moreover, systematic reviews indexed in *ACP Journal Club* must meet specific quality criteria. Thus, choosing this sampling frame allowed us to identify systematic reviews of reasonable quality (or better) that are directly relevant to clinical practice.

Cohort Selection Process

Each record identified through the search of *ACP Journal Club* was screened for eligibility on the basis of title and abstract by 2 reviewers. Records with consensus in favor of eligibility were promoted, where final confirmation of eligibility was made based on the full report. Records were screened in alphabetical order by first author until 100 eligible reviews were identified. We chose a sample size of 100 to balance the practical issue of time required to ascertain the need for updating for each review in the cohort with power considerations, such as the expected width of confidence intervals given a denominator of 100 and the ability to evaluate predictive models of the need for updating with at least 3 to 5 potential predictors in the models. Of the 100 total reviews, we set the maximum number of Cochrane reviews was to 30. We chose to limit the number of Cochrane reviews, as evidence suggests that they differ in important respects from other systematic reviews in the peer review literature on the basis of style and possibly on topic coverage.[1]

A supplemental sample was formed for question 3 as additional eligible reviews beyond the 100 had been identified, and because data extraction was quick and a larger cohort would facilitate comparisons between report types, these additional reports were included in the cohort for question 3. Few eligible HTA reports were identified through *ACP Journal Club* so Evidence Reports that were otherwise eligible were added to permit comparisons of production milestones between HTA reports undertaken by AHRQ and other types of reviews.

When an eligible review was an explicit update of an earlier review (e.g., in the case of Cochrane reviews, which are updated and reissued periodically as a matter of policy), we used the earliest version in the time frame of 1995-2005. Similarly, when more than one review on the same topic was identified, only the earliest was included, to avoid double counting the same changes in evidence (or lack thereof).

We abstracted data on primary outcomes for each systematic review. To qualify as primary outcomes, we required that authors use the words "primary" or "main" and that they identify no more than 3 such outcomes (i.e., we regarded identification of more than 3 "primary" outcomes as inconsistent with the concept of primary outcome). For reviews that did not identify primary outcomes, we selected outcomes in the order in which their results were presented, including up to 4 efficacy outcomes and up to 2 harm outcomes. Eligible outcomes were clinical outcomes (disease endpoint, functional status, mortality) or established intermediate outcome (e.g., blood pressure, glycemic control, standard instrument for measuring disease activity, such as a depression scale). Each must have provided an eligible quantitative synthesis in the formats noted above (relative risk, odds ratio, or absolute risk difference for binary outcomes and weighted mean differences for continuous outcomes).

Data Collection

The several questions reported here (detecting updating signals for the cohort of 100 quantitative systematic reviews, publication time lags for a larger cohort of 148 reviews, analysis of the patterns of growth in evidence in different clinical areas, and the survey of organizations involved in systematic review work regarding updating practices) involved different data collection methods. These details are presented in sections for each project.

I. Signals for Updating and Survival Analysis for the Cohort of 100 Systematic Reviews

Data extraction from the cohort reviews. For each of the 100 systematic reviews, we characterized the type of intervention (drug, device, or procedure), the numbers of included trials and participants, methodological features, such as the presence of heterogeneity or publication bias, descriptions of reported outcomes and identification of those explicitly identified as 'primary' or 'main,' the meta-analytic results for each outcome, and excerpted quotations of the authors' characterizations of these results and their interpretation of them.

We also classified all reviews into a clinical area. For reviews published in print journals, we primarily based this classification on the ISI classification of the clinical area of the journal in which the review appeared. For reviews published in general journals, Cochrane reviews, and HTA reports, we considered the specialty journals for which the review would have been most suitable. In the case of Cochrane reviews, we also based the classification of clinical content area on the review group that carried out the work (e.g., the Cochrane Musculoskeletal Group, the Cochrane Metabolic and Endocrine Disorders Group). For other types of reviews (e.g., HTAs), we searched the Cochrane library to find reviews on similar topics and examined the reviews to determine which review group undertook them. Two investigators undertook these classifications (AI, MS), with their results confirmed by a third reviewer (MA) with a clinical and research background.

Identification of new data for each review in the cohort. We performed systematic searches for each of the 100 reviews using a variety of electronic search strategies. Constructing searches as comprehensive as one would undertake for a formal systematic review (or an update) would involve a prohibitive amount of work given our cohort size of 100 systematic reviews.

Therefore, we adopted a combination of efficient strategies. Briefly, these involved developing simple subject searches and then limiting the results to the Core Clinical Journals subset plus the Randomized Controlled Trial publication type, subject searches run using the Clinical Query* filter in Ovid, applying the Related Articles function in PubMed® to the three largest and the three most recent studies in the original review (i.e., up to 6 studies in total), and using a 'citing references' search (through Scopus™) to identify new randomized trials that cited the original review. These search strategies served two purposes: one was to identify all new studies appropriate for updating the original systematic review; the other was to compare the performance of different strategies and evaluate their relative efficiency as surveillance methods for detecting signals for the need to update prior reviews. For studies where an updating signal occurred, we searched CENTRAL, The Cochrane Collaboration's Central Register of Controlled Trials, using the subject search developed for MEDLINE®. Examples of subject searches with the limits tested are shown in Appendix C*. A sample recording sheet used as the basis for assessing search performance is shown in Appendix D.

For each systematic review in our cohort, project team members who had backgrounds in both medicine and research screened citations retrieved by the above methods to identify trials that would have met the inclusion criteria in the original meta-analysis. Retrieved records were screened in chronological order, and the full text of articles was used when necessary to determine eligibility or extract data. The review protocol stopped when one of the signals for the need for updating (defined below) was met. Wherever possible we identified new systematic reviews on the same topic. When the search strategies yielded no eligible new trials, we conducted more comprehensive electronic searches and reviewed relevant chapters in sources such as *Clinical Evidence* and *UpToDate* to ensure that we had not missed new sources of evidence. Figure 1 outlines the overall review protocol for assessing the presence or absence of signals for updating for each of the systematic reviews in the cohort.

Outcomes: Signals for Changes in Evidence That Would Warrant Updating. Ideally, assessments of the need to update previous systematic reviews would involve assessments by experts of new evidence relevant to the original review. Shekelle and colleagues used such an approach in order to determine if guidelines required updating.[7] By choosing a small number of guidelines (17) produced by a single agency, they were able to ask the authors of the original guidelines to assess changes in evidence. This approach would clearly not be feasible for a larger sample (100 systematic reviews in the present case). It is also worth noting that identifying experts is not a straightforward task, requiring a balance of context expertise, methodological expertise, and freedom from bias regarding the question under consideration (not always easy to find among experts in a given area).

In designing a method for detecting changes in evidence without resorting to consulting experts, we considered the work of previous investigators[12-14] who have addressed similar problems involving the comparison of two sets of results related to the question—randomized and non-randomized studies of the same intervention,[14] initial and subsequent trials evaluating the same therapy,[13] and conference proceedings versus full-length journal articles for the same

* In Ovid MEDLINE®, there are three clinical queries available for therapies; sensitivity, specificity and optimized. We used the optimized query.

* Appendixes cited in this report are available electronically at http://www.ahrq.gov/clinic/tp/sysrevtp.htm.

trials.[12] In all of these examples, investigators made determinations of important changes or differences between results without resorting to expert review. They achieved such

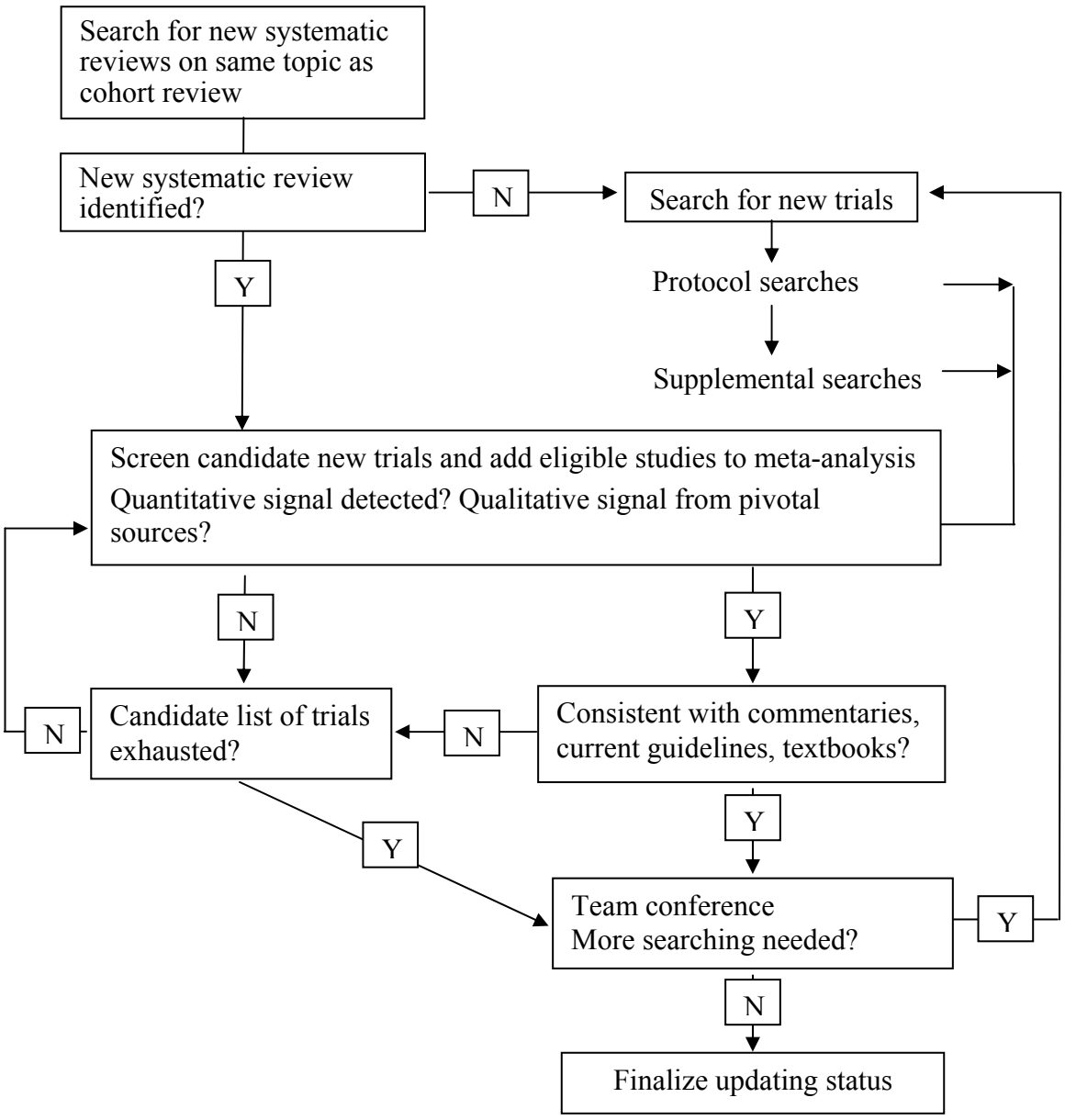


Figure 1. Review protocol to detect signals for updating.

determinations credibly by using a combination of quantitative signals (roughly the same as the ones we have chosen) and qualitative signals based on the language used to describe the results. For instance, if an article characterized a therapy as effective and another article evaluating the same therapy described it as ineffective, this would represent a major change. In a similar manner, we conceptualized quantitative and qualitative signals of potential changes in evidence sufficiently important to warrant updating of a previous systematic review.

Quantitative signals consisted of changes in statistical significance (using the conventional alpha of 0.05) or large changes in effect size (a relative change in effect magnitude of at least 50%). We restricted these changes to those involving one of the primary outcomes of the original systematic review or any mortality outcome (i.e., all-cause mortality or any cause-specific mortality outcome for which the original review provided a meta-analytic estimate of effect). We also discounted 'borderline' changes in statistical significance, which we defined as having occurred when the original and updated meta-analytic results both had p-values in the range of 0.04 and 0.06. For instance, a change from p =0.041 to p =0.059 would not count as a quantitative signal to update, nor would the converse change (from p =0.059 to p =0.041). We discounted such changes, as well as changes in effect magnitude less than 50% and all changes involving non-primary outcomes, so that quantitative signals of changes in evidence would represent robust indicators of the need to update previous reviews. Quantitative signals were detected by performing updated meta-analyses that combined data from eligible new trials with the previous meta-analytic results.

Qualitative signals of the need to update involved factors relevant to the application of evidence beyond changes in the original meta-analytic estimates. These included new information about harm sufficient to impact clinical decision making, important caveats to the original results, emergence of a superior alternate therapy, and important changes in certainty or direction of effect. Qualitative signals were detected using explicit criteria for comparing the language used to characterize findings in the original systematic review with descriptions of findings in new systematic reviews that addressed the same topic, new 'pivotal trials', new clinical practice guidelines, or new editions of major textbooks (e.g., *UpToDate*). Pivotal trials were defined as trials that had a sample size at least three times the previous largest trial or were published in one of the 5 top general medical journals (*New England Journal of Medicine, The Lancet, JAMA, Annals of Internal Medicine,* and *BMJ*) based on a ranking by journal impact factor. We defined qualitative signals with two levels of importance: signals of 'potentially invalidating changes in evidence', which we considered as changes such that one would no longer want clinicians or policy makers to base decisions on the original systematic review (e.g., a pivotal trial characterizes treatment effectiveness in opposite terms to those in the original review); and signals of 'major changes in evidence', which we regarded as changes that would not completely invalidate the previous results but would still affect clinical decision making in important ways. Such changes might include information about the way the treatment must be delivered to confer benefit, identification of populations of patients for whom treatment is more or less beneficial, or information about impact on harder outcomes than those reported in the previous systematic review (e.g., the previous review analyzed intermediate endpoints, such as blood pressure or lipid levels, whereas new trials provide data on disease end-points, such as myocardial infarction or stroke, functional status, mortality).

Major changes also included changes in characterizations of effectiveness that were less extreme than those for potentially invalidating signals, but which would still affect clinical decisionmaking. For example, whereas a change from 'effective' to 'ineffective' would represent a signal for a potentially invalidating change in evidence, a change from 'possibly beneficial' to 'definitely beneficial' would represent a major change. Importantly, no attempt was made to distinguish between varying descriptions of "possibly effective." Characterizations such as "may be effective," "promising," "trends towards effectiveness," and other similar phrases or concepts were all categorized as "possibly effective." Thus, qualitative signals for changes in evidence

captured substantive differences in the characterization of treatment effects, not merely semantic differences.

Detailed definitions of the criteria for qualitative and quantitative signals are provided in Appendix A*; Appendix B provides specific examples of their application.

Detection of Quantitative Signals for Updating

An Excel worksheet was developed in which, for a given systematic review, project team members could enter the original meta-analytic result for each outcome into a template for the appropriate format (relative risk, odds ratio, risk difference, or weighted mean difference) and enter the results of new trials identified as eligible for inclusion. For a given outcome, the work sheet allowed entry of a summary estimate or raw data (e.g., a relative risk and 95% confidence interval or the number of events for patients in each study group using the format of a two by two table).

The worksheet was programmed to perform updated meta-analytic estimates and to apply logical tests to indicate when the updated result met one of the criteria for a quantitative signal (change in statistical significance or relative change in effect size of at least 50%). Because many of the original systematic reviews included a large number of trials and often did not report data for the individual trials in a complete fashion, it was impractical for us to obtain data for each trial included in the original meta-analytic estimate. Consequently, we performed the updated meta-analyses by combining the original pooled result with the individual results of eligible new trials. With fixed effects models for meta-analysis, this procedure gives the same result as would be obtained using the individual trials from the original meta-analysis Therefore, for pragmatic reasons (avoiding having to obtain original data from each trial included in each of the100 systematic reviews) we employed fixed effects models in our updated meta-analyses. Though random effects models are usually preferred to avoid spurious precision in the face of heterogeneity, we regarded this approach as reasonable, since our goal consisted of detecting changes in evidence that had likely occurred, not producing exact estimates of updated treatment effects.

Data from new trials were entered into the meta-analytic calculator in chronological order, so that the time at which a quantitative signal was met could be identified. In general, we stopped the review protocol once a change in statistical significance or change in effect size of at least 50% occurred, though we sometimes continued to add new trials to confirm stability of the results.

Group review and classification. After assessment by the reviewer, each systematic review in the cohort was discussed at a case conference attended by the team of KS, MS, MA, and JJ. At this meeting, the final classification of the updating signal status was decided by group consensus, and the completeness of the evidence base was discussed. The team had the option to request additional searching, or search directly for new studies known or suspected by team members to be relevant.

Date definitions for survival analysis. Two survival analyses were undertaken. The first used the publication date as birth. We used the MEDLINE Entrez date as a surrogate for

* Appendixes cited in this report are available electronically at http://www.ahrq.gov/clinic/tp/sysrevtp.htm.

publication date of the systematic review, as this date always includes a day, month, and year (not always the case for journal publication dates) and because the Entrez Date closely follows the publication date (typically within days to several weeks). In a second survival analysis, we defined birth as the end of the search period reported in the review. (This date did not always include a day and month. We imputed all missing months as June and all missing days as the 15th.) The end point, 'death' for both survival analyses was the Entrez date associated with the new evidence that resulted in the signal for updating. Where the updating signal derived from non-MEDLINE sources (e.g., an advisory from the Centers for Disease Control and Prevention, the Food and Drug Administration, or a chapter in a textbook), we used the date of publication as the date of the signal for updating. For surviving systematic reviews, observations were censored on September 1, 2006, the approximate midpoint of the 4-month period during which searches were performed for the entire cohort.

Performance of the surveillance searches in detecting signaling evidence. Three main types of signaling evidence were used; new RCTs that were added to a meta-analysis from the original systematic review in the manner of a cumulative meta-analysis, single RCTs that met our criteria for a pivotal trial, and new systematic reviews that provided evidence that appeared to overturn the findings of the original review, either by contradicting the original findings, adding an important caveat or demonstrating a significant harm. Other signaling evidence (i.e., evidence that provided the basis for signals) included FDA advisories and expert opinion from *UpToDate*, and clinical trials that did not meet the criteria for pivotal trial. These sources were used as sources of signals for updating in only five reviews.

The surveillance searches looked for primary studies and for systematic reviews with the publication type meta-analysis in MEDLINE. To determine the effectiveness of these searches to detect signaling evidence, we examined recall of signaling articles in the subset of systematic reviews studies here were those updated by search (n=79), and for which a major or notable signal occurred. For the analysis of RCTs added to the cumulative meta-analysis, only those systematic reviews which also had a quantitative signal and were updated by search were studied.

Any signaling evidence added by nomination was tested to determine if it was indexed in MEDLINE and if would have been retrieved by the searches. In some cases, the evidence was published after the searches for new evidence for that systematic review were run. The database was updated with the search results for those nominated publications.

Targets for the cumulative meta-analysis were any RCT added to the meta-analysis of the outcome which had the signal, up to the point where the signal occurred. Targets for the final RCTs were the pivotal RCTs. Targets for final MAs were the newer meta-analyses that contained the evidence that rendered the cohort systematic review potentially in need of update. These were meta-analyses that were not explicit updates. Finally, all signaling evidence was considered. For each of these analyses, recall was calculated for each type of search. For the final analysis of recall of any signaling evidence, two additional variables were created representing recall from MEDLINE by any of the subject search methods (CQ, AIM RCT or MA) and recall from MEDLINE by either of the related articles search methods (RI RCT and RI MA).

II. Publication Time Lags

For question 3, the impact of publication time lags on updating, we supplemented the data set used in the survival analysis with additional eligible systematic reviews identified through *ACP Journal Club*, as well as AHRQ Evidence Reports that met all eligibility criteria for the main cohort, except inclusion in *ACP Journal Club*.

We determined dates for performance of the original search, manuscript acceptance, and publication of the review. We regarded the search date as the most recent date reported in the methods section of the systematic review. For Cochrane reviews, we used the most recent of the following dates: the search date reported in the search strategy section in the body of the review, the date new studies were found and included/excluded (e.g., for updated reviews), or the date new studies were sought but not found. For database dates, the end date reported for MEDLINE searching was used (i.e., 1966-June Week 4, 2003) if available. If the MEDLINE date was not reported, any other database end date was used. If no end date was reported, the variable was treated as missing. For all types of reviews, the publication date and indexing date was taken from the Ovid MEDLINE records.

For each date (original search, manuscript acceptance, publication), we identified a year, month, and day. When month was missing, we imputed the 6th month; when day was missing, we imputed the15th day of the month.

III. Growth of the Literature by Clinical Area

MEDLINE searches based on high-level MeSH headings corresponding to the ISI journal categories were undertaken. The resulting set of citations was limited to the publication type Randomized Controlled Trial, to the publication type Clinical Trial but not Randomized Controlled Trial, to the MEDLINE Systematic Review subset, and to the publication type Clinical Practice Guidelines. Searches were then limited by year for each year between 1988 and 2006. We chose 1988 as the beginning of the time period of interest, as this date corresponded to the period five years prior to the earliest search date for any systematic review in the cohort. Search strategies are illustrated in Appendix E*.

IV. Survey of Organizations Engaged in Funding or Production of Systematic Reviews

This exploratory Internet pilot survey on current updating practices and policies employed a purposeful sample to allow for investigation of likely information-rich cases. We chose 9 organizations well known to fund or carry out systematic reviews, as well as 12 EPCs, in addition to AHRQ, were also asked to complete this survey. The identities of the organizations have been kept anonymous per the statements contained in the informed consent signed by participants in the survey and as stipulated in the research protocol approved by the institutional ethics review board at the Children's Hospital of Eastern Ontario.

* Appendixes cited in this report are available electronically at http://www.ahrq.gov/clinic/tp/sysrevtp.htm.

The survey was provided to participants via the Survey Monkey[15] web-based service. This was considered a suitable forum given distribution of the sample across a wide international geographical area, and that key informants are frequent Internet users with email addresses.[16,17] Emails were sent directly to organizational Directors or to the highest ranking scientific or administrative official, asking them to identify the most appropriate internal respondent to answer the questionnaire. Data collection consisted of approximately 50 questions (including skip-logic functionality). These questions focused on the following topics: (a) updating policies, (b) responsibility for updating, (c) estimates of outdated reviews, (d) updating strategies and practices, including when to update, surveillance and triggers impacting updating decisions, (e) strategies for how to conduct an update, (f) barriers and facilitators to this process, (g) views on updating collaboration between groups and (h) descriptive demographics and characteristics of the organization and the representative key informant. It was estimated the survey took between 20 to 30 minutes to complete. (Appendix G*: Survey Instrument)

We attempted to increase our overall response rate by employing recommended survey methods to maximize Internet survey participation.[17-20] Participants were contacted four times. A small financial incentive was offered to all participants who completed the survey. On clicking on the link to the survey, participants were presented with a description of the purpose of the study, assurance of confidentiality, and a statement of the research protocol by the hospital ethics review board, followed by a request to provide informed consent or decline participation in the survey. Reminder emails were scheduled for day 10, 15 and 25 of the survey.

Analysis

We fit non-parametric Kaplan-Meier curves to the data set of censored and uncensored observations and used multivariable proportional-hazards models to examine the association between survival and various features of the systematic reviews at the time of publication. We distinguished two categories of potential predictors of survival. The first category consisted of features knowable at the time of publication for a given systematic review, including clinical content area (e.g., cardiovascular medicine, obstetrics and gynecology, critical care, infectious diseases), numbers of participants and trials included in the meta-analysis, the identification of heterogeneity or publication bias, 'recent or ongoing activity in the field', which we defined as present if the review included at least one trial published within the last year of its search period or if the review identified ongoing trials eligible for inclusion. Because some evidence exists to suggest that Cochrane reviews differ in important ways from other systematic reviews,[1] we also included a dichotomous variable for Cochrane review versus other systematic reviews. The second category of predictors consisted of features knowable only after some surveillance of the literature (but not performance of a full update of the review). Such predictors included the number of new trials eligible for inclusion in an update of the original review, the number of new participants in these trials, the ratio of the new total number of trials to the previous total, and the ratio of the new total number of participants to the previous total.

After confirming that the assumption of proportionality applied, we performed stepwise multivariate analyses using a threshold of $p \leq 0.1$ for variable selection and retention. In addition to the proportional hazards analysis to estimate predictors of survival, we conducted logistic

* Appendixes cited in this report are available electronically at http://www.ahrq.gov/clinic/tp/sysrevtp.htm.

regression analysis to identify predictors of survival less than two years. Cohort members that were censored in less than two years were counted as missing for this analysis. All analyses were performed with SAS version 9.0 (The SAS Institute, Cary, North Carolina).

Analysis of group differences in time lags in the publication process was made using nonparametric statistics (e.g., Kruskal-Wallis test for differences in median publication times between groups).

Survey Analysis

Closed-ended questions were analyzed primarily using a descriptive summary of findings in the form of frequencies. In addition, percentages were calculated and other details reported in text and tabular form. Participating organizations were not identified in the results as only aggregate data is reported. The EPCs also responded to several additional open-ended questions of particular interest to the AHRQ's EPC Program. The responses to these supplemental questions were compiled for internal use by the AHRQ and are therefore not discussed in this report.

Chapter 3. Results

Results of Literature Search and Cohort Screening

Records for 651 potentially eligible systematic reviews were identified through searching. Achieving our target sample size of 100 reviews for the analysis of updating signals required that we assess a total of 325 reviews for eligibility. We screened additional reviews to add a further 50 reviews to the set of reviews in the analysis of publication time lags. (The analysis of time lags was less labor intensive, permitting a larger cohort size for this part of the project). 165 records were excluded on the basis of the *ACP Journal Club* record, and 60 articles were excluded after assessment of the full article. Exclusion reasons are shown in Figure 2.

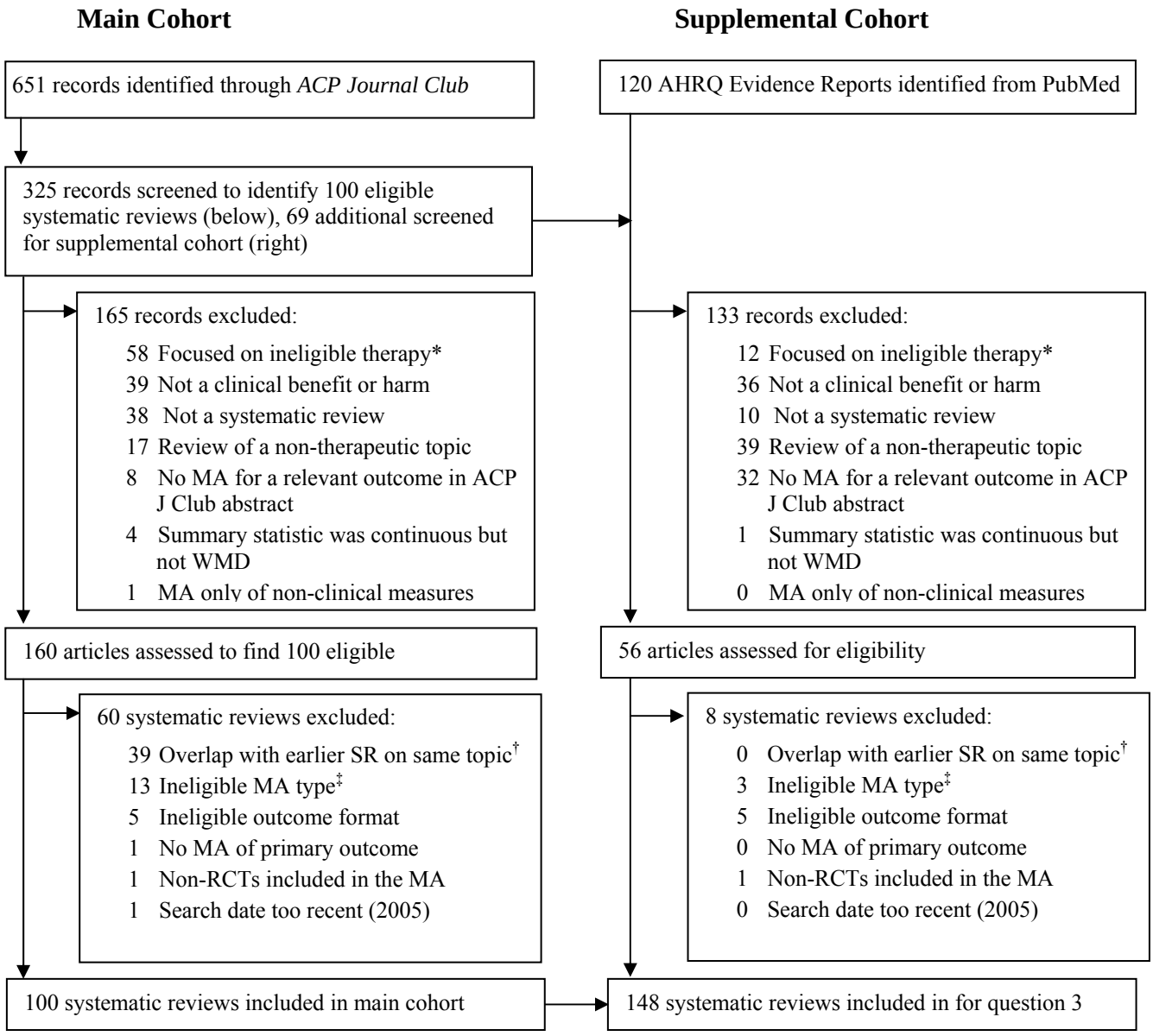


Figure 2. Flow of information through eligibility assessment.

*This category includes reviews not focused on a specific class of drug, device or procedure, as well as ones focused on educational or behavioral interventions, or complementary therapies.

†This category includes updates of systematic reviews already in cohort, topics similar to that of a systematic review already included, or the journal version of an included Cochrane review

‡ This category includes meta-analysis using individual patient data without regular meta-analysis, meta-regression, or indirect meta-analysis.

Assessment of the New Evidence

Seventy-seven of the systematic reviews were assessed against new evidence found (if any) through searching and 23 were assessed against an updated systematic review. The updated review could be either an update performed by the authors of the original review, or a newer review on the same topic identified through the search of *ACP Journal Club* that would itself be eligible for inclusion in the cohort.

Characteristics of Included Studies

Composition of the Cohort

Each review in the cohort of 100 systematic reviews included a median of 13 studies (inter-quartile range: 8 to 21) and 2663 participants (inter-quartile range: 1281 to 8371). We were able to identify at least one new eligible trial for 85 systematic reviews, with a median of 4 new trials (inter-quartile range: 1 to 7) and 1160 new participants (inter-quartile range: 170 to 3689) per review. The five most common clinical content areas for the original systematic reviews were cardiovascular medicine (20), gastroenterology (13), neurology (11), infectious diseases (9), and respiratory system (9). Only 15 of the reviews evaluated the effects of medical devices or procedures; drug therapies provided the focus for the rest of the cohort (Table 1).

Table 1. Characteristics of the Cohort of 100 Systematic Reviews

Characteristic	Composition of Cohort
Publication Type	Peer-reviewed journal article (72), Cochrane review (27), Other* (1)
Intervention Type	Medication (85), Medical device (8), Procedure (7)
Clinical Categories	Cardiovascular (20), Gastroenterology (13), Neurology (11); Other 10 categories each included fewer than 10 systematic reviews
Publication Period	
January 1995 to February 28, 1997	16
March 1997 to April 30, 1999	22
May 1999 to June 30, 2001	25
July 2001 to September 30, 2003	21
October 2003 to December 31, 2005	16
Dates of searches	
June 1990 to April 1993	2
May 1993 to March 1996	21
April 1996 to February 1999	28
March 1999 to January 2001	30
February 2002 to December 2004	18
Search date not reported	1
Source	The Cochrane Library (27), BMJ (19), JAMA (7), Lancet (7), Annals of Internal Medicine (6), Archives of Internal Medicine (5) and 25 other titles
Median Number Included Trials	13 (inter-quartile range: 8-21)
Included Number Included Participants	2663 (inter-quartile range: 1281-8371)
Included ≥ 1 trial published within last year of search period	67 reviews
Original review mentioned ongoing trials	26 reviews
Heterogeneity as assessed by authors of original review	Identified as statistically significant: 50 Not statistically significant but authors still suspected as possibly present: 11 Heterogeneity regarded by authors as absent: 32 Not assessed: 7
Publication bias as assessed by authors of original review	Identified as statistically significant: 4 Not statistically significant but authors still suspected as possibly present: 14 Regarded by authors as absent: 22 Not assessed: 60

* Other: 1 systematic review published by the Canadian Coordinating Office for Health Technology Assessment

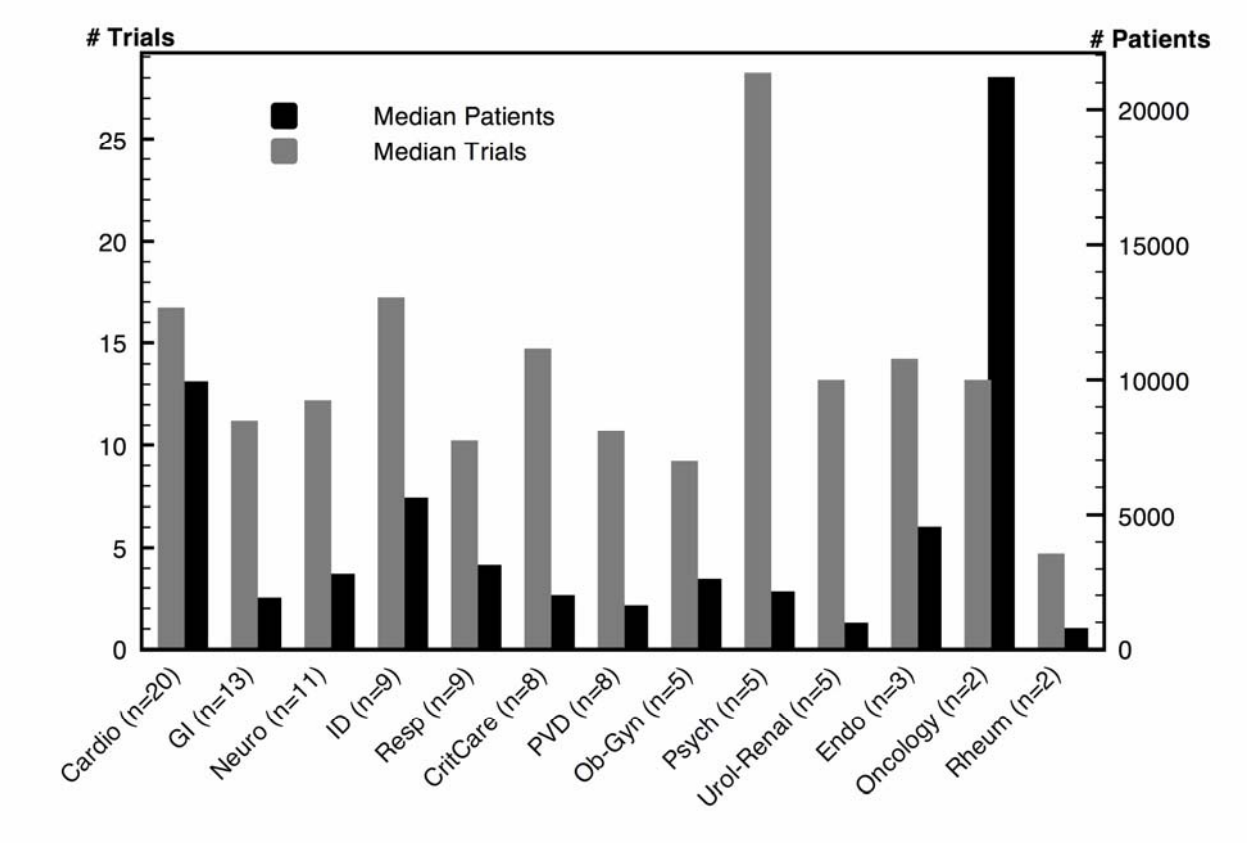


Figure 3. Median number of trials and median number of trial participants included in systematic reviews by clinical area.

Parenthetic numbers beside labels on horizontal axis indicate the numbers of reviews in the cohort. The horizontal axis is arranged in descending order of frequency by clinical topic area, with cardiovascular medicine (20 systematic reviews) and gastrointestinal diseases (13 systematic reviews) at the far left and oncology and rheumatology at the far right (2 systematic reviews in each category).

Cardio=Cardiovascular, GI= Gastrointestinal diseases, Neuro=Neurology, ID= Infectious Diseases, Resp = Respiratory diseases, CritCare=Critical Care, PVD= Peripheral Vascular Diseases, Ob-Gyn= Obstetrics and Gynecology, Psych= Psychiatry, Urol-Renal= Urology & Nephrology, Endo=Endocrinology and metabolism, Rheum=Rheumatology

Signals for Updating

Of the 100 systematic reviews, a quantitative signal for updating occurred in 30 cases. Qualitative signals for the need to update occurred in 54 cases, including 8 that met criteria for a potentially invalidating change in evidence and 46 that met criteria for a major change. Qualitative signals had their basis in new systematic reviews in 23 cases (including explicit updates in 5), pivotal trials in 25 cases, and other sources in 6 cases (trials discussed in *ACP Journal Club* or *UpToDate* and advisory statements issued by the Centers for Disease Control and Prevention or the Food and Drug Administration). The primary event of interest, a quantitative signal involving the primary outcome of the original systematic review or qualitative

signal for potentially invalidating or major changes in evidence, occurred for 57 reviews (57%; 95% CI: 47% to 67%) in the cohort (

Table 2).

Table 2. Frequency of the Different Types of Signals for Updating

Type of Signal for Updating	Number of systematic reviews in cohort
Quantitative signal	20
Change in statistical significance	18
Relative change in effect size ≥ 50%	12
Qualitative signal	54
Opposing findings	7
Substantive changes short of opposition	16
Clinically important caveats	28
Clinically important expansion of therapy	3
Harm that completely undermines therapy	1
Superior alternate therapy	1
Primary event of interest: either Qualitative or Quantitative signal	57

Survival Analysis

Using publication date as 'birth', median event-free survival (i.e., time without a signal for updating) was 5.5 years (95% CI: 4.6-7.6). However, in 23 cases, signals for updating occurred in less than 2 years, and in 15 cases the signal occurred in less than 1 year. In 7 cases, a signal had already occurred at the time of publication of the original systematic review (in one case, 295 days prior to publication).

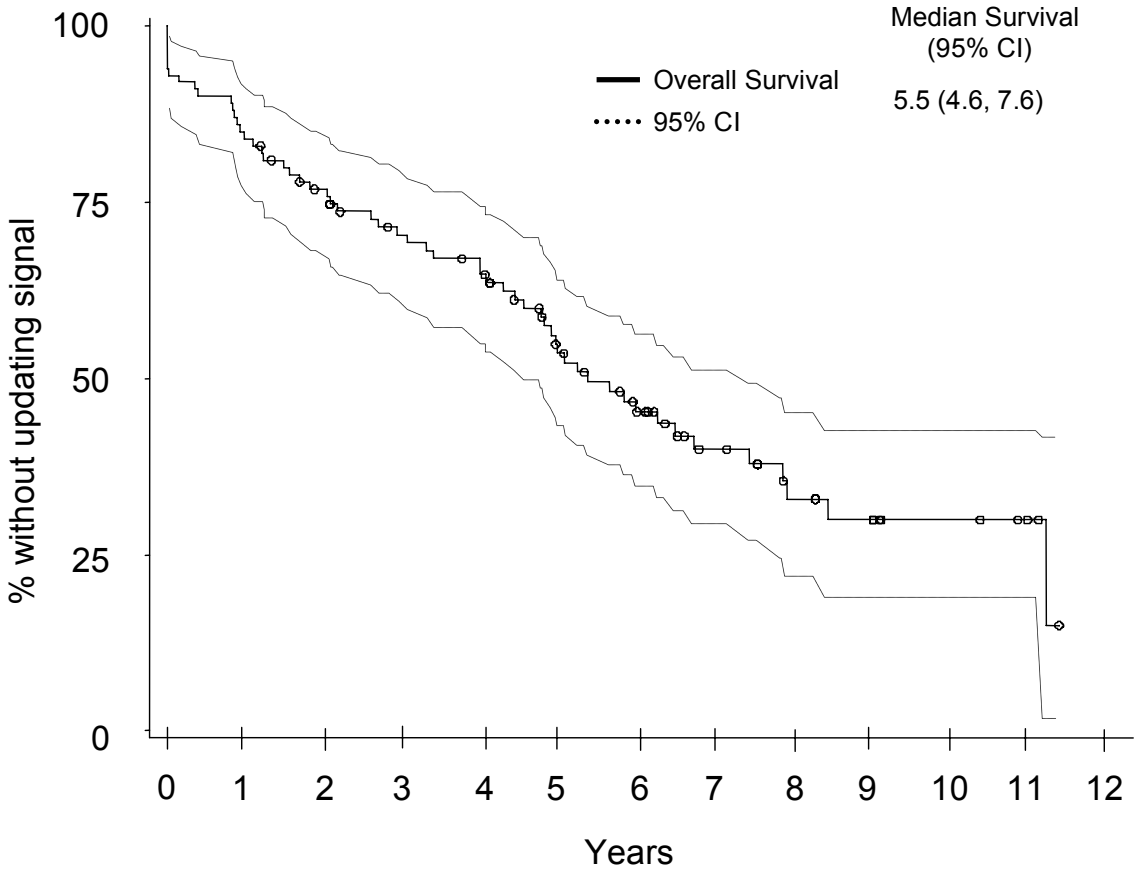


Figure 4. Kaplan Meier plot showing the overall event free survival (time without a signal for updating) using publication date as 'birth'; the immediate drop in survival at time zero reflects the 7 systematic reviews for which signals for updating had already occurred at the time of publication. Symbols represent censored cases.

In univariate analyses, shorter survival was associated with a clinical content area of cardiovascular medicine (hazard ratio [HR] of 2.58, 95% CI: 1.39 to 4.78; p =0.003), increase in the total number of patients by a factor of 2 or more (HR of 1.79; 95% CI: 1.03 to 3.10; p =0.04), and heterogeneity for at least one outcome in the original systematic review (HR of 1.64, 95% CI: 0.94 to 2.86;p =0.08) (Figures 4-7). Other potential predictors evaluated, but not found to significantly affect survival included: the number of included patients in the original review, identification of publication bias in the original review, the inclusion of at least one trial published in final 12 months of the search period, the identification of ongoing trials, and the publication type (Cochrane reviews versus those published in peer review journal articles) (Table 3).

Table 3. Univariate Survival Analysis

	Hazard Ratio	P-value
Clinical Category		
Neurology	1.37 (0.59, 3.16)	0.47
Cardiovascular	2.58 (1.39, 4.78)	0.003
Gastroenterology	1.35 (0.58, 3.13)	0.48
Other	reference	-
Heterogeneity present or suspected	1.64 (0.94, 2.86)	0.08
Publication bias present or suspected	0.99 (0.46, 2.12)	0.98
Activity in field *	1.36 (0.76, 2.44)	0.30
Number of included studies > median (13) for cohort	0.79 (0.46, 1.33)	0.37
Number of included participants > median (2663) for cohort	1.22 (0.72, 2.06)	0.47
Ratio of New to Original Total N > 2[†]	1.79 (1.03, 3.1)	0.04
Any of 3 criteria for substantial increases in number of new trials or patients[‡]	1.17 (0.59, 2.32)	0.66
Cochrane review	0.74 (0.39, 1.39)	0.35
Largest new trial larger than previous largest N trial	1.04 (0.99, 1.1)	0.09

*Recent activity defined as present if original systematic review included at least 1 trial published within the final 12 months of the search period or if original systematic review identified ongoing trials eligible for inclusion. This variable was coded as present for 71 of the included systematic reviews.

[†] Ratio of New to Original Total N > 2 (i.e., increase in total sample size by more than a factor of 2)

[‡] Size criteria C1-C3 defined as any of the following occurring: increase in total number of trials by ≥ 50%, increase in total number of participants by ≥ 50%, publication of a new trial with sample size ≥ 3 times size of previous largest trial

Table 4 shows the results of multivariate the analysis with adjustment for all variables shown. We also performed stepwise multivariate analysis using a threshold of $p \leq 0.1$ for variable selection and retention, which resulted in a model in which the following variables predicted decreased survival: clinical content area of cardiovascular medicine (HR of 3.26, 95% CI: 1.71 to 6.21; p =0.0003), heterogeneity in the original systematic review (HR of 2.23, 95% CI: 1.22 to 4.09; p =0.01), and the ratio of the largest new trial to the largest trial from the original review (HR of 1.08, 95% CI: 1.02 to 1.15; p =0.01). Systematic reviews with more than the median of 13 included studies had increased survival (HR of 0.55; 95% CI: 0.31 to 0.98; p =0.04).

In logistic regression analysis no variable significantly affected the risk of a signal for updating occurring within 2 years of publication, though trends towards increased risk were observed for cardiovascular topics (odds ratio of 2.67; 95% CI: 0.88 to 8.1, p =0.08) and an increase in the total number of patients by at least factor of 2 (odds ratio of 2.29; 95% CI: 0.84 to 6.25, p =0.11). A trend towards decreased risk of a signal for updating occurring within 2 years was seen for systematic reviews with more than the median of 13 included studies (odds ratio of 0.38; 95% CI: 0.14 to 1.04; p =0.06). Varying the time period of interest (e.g., predicting a signal for updating within 1 year or 3 years of publication) did not substantially alter the results.

Table 4. Multivariate Analysis of Hazards

	Hazard Ratio	**P-value**
Clinical Category		
Neurology	1.38 (0.52, 3.70)	0.52
Cardiovascular	3.09 (1.47, 6.52)	0.003
Gastroenterology	1.44 (0.57, 3.62)	0.44
Other	reference	-
Heterogeneity present or suspected	2.22 (1.21, 4.08)	0.01
Publication bias present or suspected	1.06 (0.47, 2.41)	0.89
Activity in field *	1.45 (0.74, 2.82)	0.28
Number of included studies > median (13) for cohort	0.42 (0.21, 0.81)	0.01
Number of included participants > median (2663) for cohort	1.56 (0.79, 3.08)	0.20
Ratio of New to Original Total N > 2[†]	1.86 (0.95, 3.61)	0.07
Any of 3 criteria for substantial increases in number of new trials or patients[‡]	0.97 (0.45, 2.12)	0.94
Cochrane review	1.35 (0.62, 2.97)	0.45
Largest new trial larger than previous largest N trial	1.06 (0.99, 1.14)	0.12

*Recent activity defined as present if original systematic review included at least 1 trial published within the final 12 months of the search period or if original systematic review identified ongoing trials eligible for inclusion.

[†] Ratio of New to Original Total N > 2 (i.e., increase in total sample size by more than a factor of 2)

[‡] Size criteria C1-C3 defined as any of the following occurring: increase in total number of trials by ≥ 50%, increase in total number of participants by ≥ 50%, publication of a new trial with sample size ≥ 3 times size of previous largest trial

Survival contrasting the significant predictors with the rest of the cohort is illustrated in Figures 5 through 7.

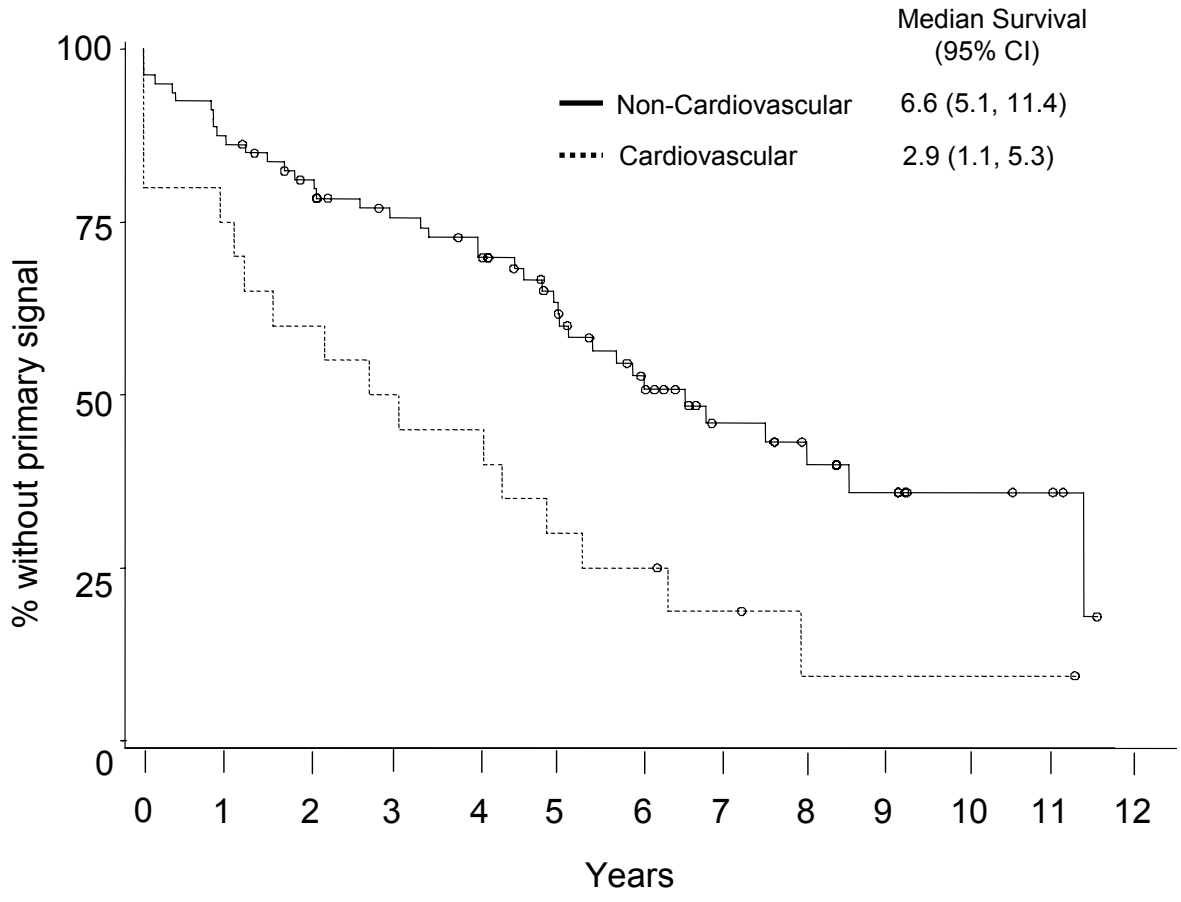


Figure 5. Kaplan Meier plot showing survival by clinical topic area of the original systematic review, stratified by cardiovascular (n=20 reviews) versus all other topics (n=80). Symbols represent censored cases.

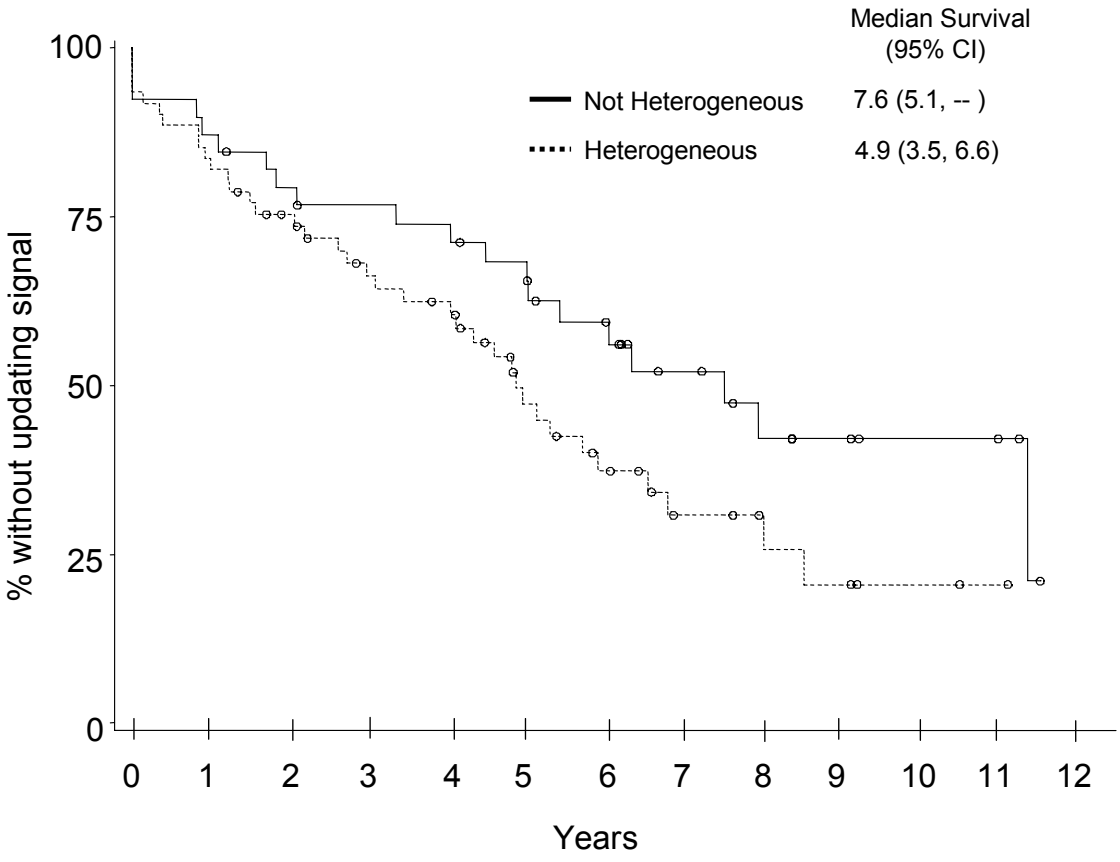


Figure 6. Kaplan Meier plot showing survival stratified by the presence or absence of heterogeneity in the systematic review; statistical heterogeneity was identified as definitely or likely present for at least one outcome in 61 of the 100 reviews. Symbols represent censored cases.

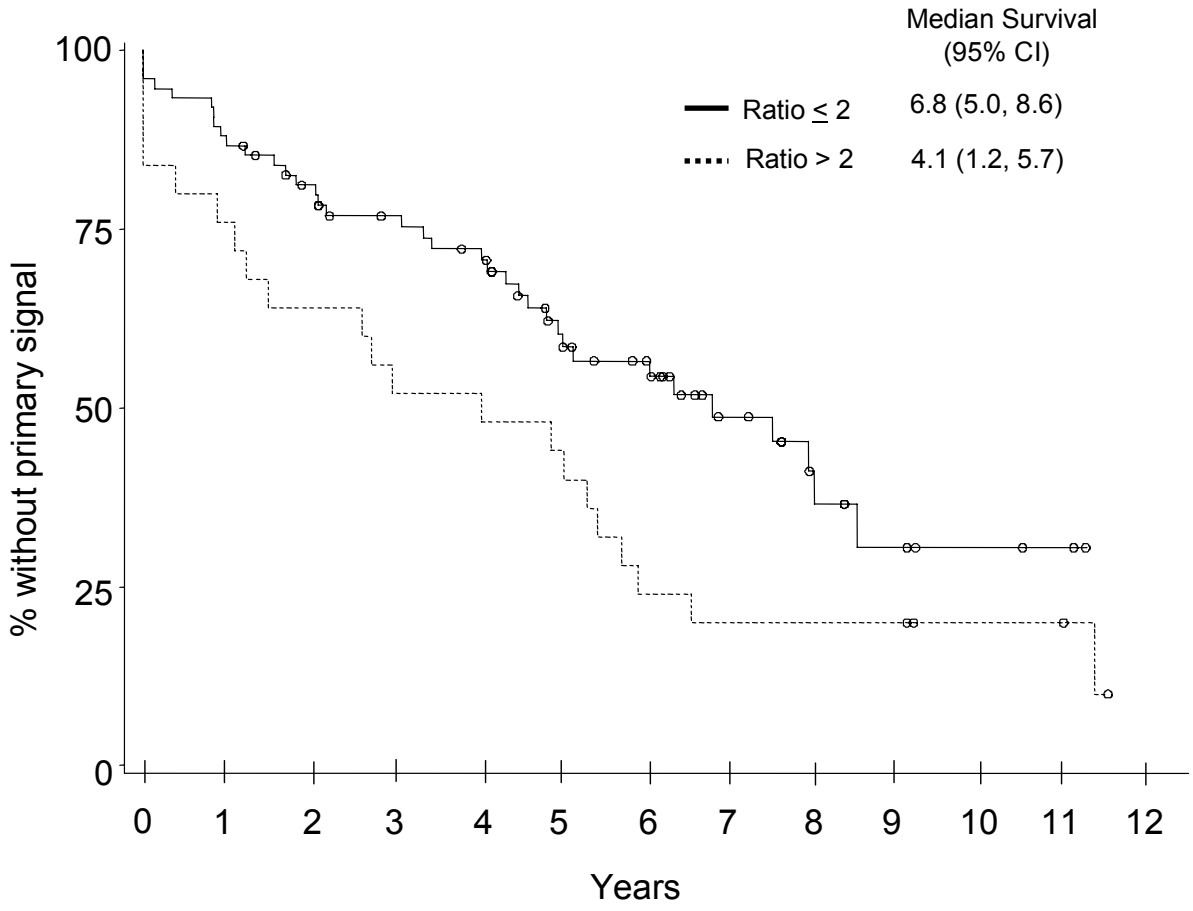


Figure 7. Kaplan Meier plot showing the effect on survival of a doubling of the total number of patients (i.e., ratio of new total sample size to old total > 2), which occurred for 25% of systematic reviews in the cohort. Symbols represent censored cases.

When survival analyses were repeated using the end of the search period as 'birth', rather than the publication date, the median survival was 6.9 years (95% CI: 6.1 to 9.0), with a median time to a signal for updating of 4.3 years (inter-quartile range: 2.1- 6.4 years). The signal for updating occurred within 1 year of the search in 4 cases, within 2 years of the search in 11 cases and within 3 years of the search in 20. Predictors of increased or decreased survival did not differ from those identified in the analysis that used publication date as 'birth.'

Directions of Changes in Evidence and Expected Impact on Practice

Of the 18 reviews with changes in statistical significance, 13 involved a gain of statistical significance (i.e., a previously non-significant result became statistically significant) and 5 involved a loss of significance. For the 12 reviews with a relative change in effect size of at least 50%, 3 involved an increase in effect magnitude and 9 involved a decrease in effect. However,

because these outcomes could involve harms or benefits, we also characterized the expected impact on practice of the changes that gave rise to the signal for updating. Increases in magnitude of benefit, certainty about benefit, or identification of new patient populations that benefit from the treatment were classified as expected increases in therapeutic application. Decreases in magnitude of benefit, decreased certainty about benefit, findings of increased harm or other limitations on benefit were all classified as leading to decreased therapeutic application. Using such explicit criteria, use of the therapies evaluated would be expected to increase in 19 and decrease in 28 (Table 5).

We also assessed the impacts on certainty of results due to the changes in evidence that gave rise to signals for updating. We characterized changes in certainty using the 5-point scale that formed the basis for judging characterizations of effectiveness (Appendix A*). This scale included the following categories: definitely effective, probably or possibly effective, uncertain effectiveness, probably or possibly ineffective, and definitely ineffective. When the updated result lay further from the middle position (complete uncertainty) than the original result, we regarded certainty as having increased. Conversely, when the updated result lay closer to the middle position than the original result, we regarded certainty as having decreased. When the updated and original results were equally distant from the middle position (e.g., definitely effective and definitely ineffective), we did not regard certainty as having changed. Such cases would, however, count as impacting therapeutic use. As shown in Table 5, the majority of signals for updating involved increases in certainty (30 reviews) or no changes in certainty (50 reviews).

Table 5. Changes in Certainty and Expected Impacts on Practice Associated with Signals for Updating.

Impacts of new evidence on certainty of results		Expected impacts of new evidence on clinical practice	
Increase in certainty	30	Increased therapeutic use	19
Decrease in certainty	3	Decreased therapeutic use	28
Unchanged certainty	50	Unchanged therapeutic use	44
Unclear change	17	Unclear change	9

Search Performance

Across all reviews, 477 new reports were identified as eligible for inclusion to the systematic reviews. Of these, searching identified 430, and 47 were identified by the reviewers from among the studies included in meta-analysis retrieved by the subject search or related item searches. Forty of these nominations (85%) were indexed in MEDLINE, thus the searches retrieved 92% of eligible new studies identified. Forty-three of the 47 missed studies were from systematic reviews where we had searched CENTRAL. Two of these 43 nominations were retrieved by the CENTRAL search.

* Appendixes cited in this report are available electronically at http://www.ahrq.gov/clinic/tp/sysrevtp.htm.

In 59 cases, a single search strategy would have been sufficient to retrieve all eligible new studies found by any method. Related article RCTs was sufficient in 45 cases, Clinical Query in 34, Core Clinical Journal RCTs in 9, CENTRAL in 14 and Citing RCTs in 3. Systematic reviews with multiple sufficient strategies tended to be those with few new studies. In 68 cases, searching Related Article RCT and Clinical Query would have retrieved all studies either because one strategy or the other was sufficient or because the two together was sufficient.

The median number of records retrieved by the combination of Related Article RCT and Clinical Query by the date at which the signal for updating was detected was 71 (1st and 3rd quartiles; 25, 106). The median number of records retrieved by this combination and assessed as on topic was 7 (1st and 3rd quartiles; 4, 24).

To identify newer quantitative systematic reviews, the Related Article search and the subject search were limited to publication type of meta-analysis. The Related Article search recalled 45% of the meta-analyses found to be on topic and the subject search limited to the meta-analysis publication type identified 66% of the on topic meta-analyses. Of the records assessed by the review team, precision (proportion of assessed records found to be relevant) of the subject search was 0.38 and precision of the Related Article MA search was 0.36.

Performance of the surveillance searches in detecting signaling evidence. There were 27 final RCTs in cohort systematic reviews that were updated by searching and had a qualitative signal of major or notable. Sixteen of these also had a quantitative signal and so formed the basis of the analysis of success in detecting RCTs added to the cumulative meta-analysis.

Six of the 27 final RCTs were by nomination and the remaining 21 were found by the search. Three of the nominations were recent, high profile trials. These were used rather than reviewing the candidate list, thus for the purpose of evaluating search performance. As these occurred after the search date, these three were tested to see if the search would have retrieved them, and whether they cited the cohort systematic review. The remaining 3 nominated final RCTs were identified through meta-analysis. There were 34 targets for studies added to cumulative meta-analysis; 27 were candidates found through searching, 5 were nominations, 2 were meta-analyses found through searching where the individual trial data could not be extracted. One of the nominations was a trial published after the search date, and was manually tested to see which searches would have retrieved it. There were 9 signaling meta-analyses – one was nominated, all others were identified through searching.

Other signaling evidence was used in only 5 reviews that were updated by search. Evidence included FDA advisories and expert opinion from *UpToDate*, and clinical trials that did not meet the criteria for pivotal trial. Three of these 5 sources were indexed in MEDLINE. Two were identified through searching.

Recall by each search of each type of evidence is shown in Table 6. Retrieval was best for RCT and MA evidence, but the search methods did retrieve some of the other evidence. Overall search performance of final evidence stood at 0.65 for subject search methods, 0.76 for related article methods, 0.55 for CENTRAL and 0.17 for citing references. Across all applications in which the citing reference technique was tested, its strongest performance was in detecting other final evidence, with 0.33 recall. One of the highest recall scores seen in this study was recall of 0.89 for final RCTs found through related article searching. In general, search methods showed somewhat higher recall for final evidence than for all evidence found relevant to the reviews, and the relative performance of the various methods was similar to that seen in the more general context.

Most information was found through searches. Of 62 pieces contributing to the signal, 57 (92%) were identified through the searches of MEDLINE. The additional material was an included study in a systematic review identified through searching, known to the team, or *UpToDate*.

Table 6. Recall of Signaling Evidence by the Surveillance Searches

	In Quantitative Signal n=34	Final RCT N=27	Final MA N=9	Other Final Evidence N=5	Any Signal N=62
Related Article search with RCT limit	0.74	0.89	0.00	0.20	0.61
Subject search limited to Core Clinical journals and RCT publication type	0.41	0.67	0.11	0.20	0.40
Citing Reference search	0.18	0.22	0.00	0.33	0.17
Subject search with Clinical Query limit	0.56	0.67	0.44	0.40	0.55
Subject search with meta analysis publication type limit	0.03	0.00	0.44	0.00	0.08
Related Article search with meta-analysis limit	0.06	0.00	0.78	0.00	0.15
Retrieved by any subject search method	0.68	0.70	0.67	0.40	0.65
Retrieved by any related article method	0.79	0.89	0.78	0.20	0.76
Indexed in MEDLINE	33/34 (0.97)	27/27 (1.00)	9/9 (1.00)	3/5 (0.60)	56/59 (0.95)
Found by any search	0.94	1.00	1.00	0.40	0.92

Adequacy of MEDLINE coverage for surveillance. While there is general agreement that searching a single database is inadequate for developing the evidence base for systematic reviews,[21] the adequacy of MEDLINE for detecting the need to update (surveillance searching) has not been previously examined. We consider the proportion of studies in the original reviews that were indexed in MEDLINE, the survival of those in known updates from this sample, and the proportion of new relevant studies identified from any source that were indexed in MEDLINE.

Original systematic reviews: Of 2065 reports included in the original systematic reviews, 407 (25%) were not indexed in MEDLINE. MEDLINE indexed publications accounted for 89% of total number of participants (N) included in the original systematic reviews, although we could not identify values for N in all cases, and 40% of cases with missing N were for non-indexed studies. For reports where we could identify N, the median size was larger for MEDLINE indexed studies compared with non-indexed studies (116 participants [inter-quartile range: 43-365] vs. 80 participants [40-224]).

Updated Cochrane reviews. Of original Cochrane reviews assessed through an update, 95 studies were indexed in MEDLINE. Of these 95 studies, 4 (4%) were excluded by the author in the update. Among the 56 studies not indexed in MEDLINE, 13 (23%) were excluded in the update (odds ratio 0.145, CI 0.045-0.472), suggesting that material from sources not indexed in MEDLINE may become less important over time.

New studies. The indexing status and number of new studies assessed as eligible for inclusion in the reviews were considered. New studies included candidates identified through searching, nominations found through newer meta-analyses or known to our team, and studies included in explicit updates. Of 590 studies assessed as eligible, 33 (6%) were not indexed in MEDLINE. These reports accounted for 5503 of 648531 new participants (N) identified (1%). All pivotal trials, those RCTs that, by themselves, provided in signal for update, were indexed in MEDLINE (n=19).

Time Lags in the Production and Publication of Systematic Reviews

One hundred and forty-eight reports were included in this analysis, of which 91 (62%) were journal published reviews, 36 (24%) were Cochrane reviews and 21 (14%) were HTA reports. Of HTA reports, 19 (90%) were AHRQ evidence reports. For Cochrane reviews, we used the most recently published version of the Cochrane review.

The median time from last reported search date to indexing was 75 weeks with an inter-quartile range of 52 to 111 weeks. Lag from last search date to publication is shortest for Cochrane reviews (median 31 weeks, inter-quartile range: 22-65) and longest for journal reviews (median 69 weeks; inter-quartile range: 44-92), with technical reports falling in between (median 58 weeks; inter-quartile range: 45-74) (Kruskal Wallis χ^2 11.24, p =0.004) For reviews assessed for need of update, 7 were found to have gone out of date by the time of publication.

Intermediate milestones of submission and acceptance dates were reported only for journal published reviews, but reveal what proportion of total preparation time is under the control of investigators. For journal-published reviews where submission and publication dates are known (n=17) median processing time was 41 weeks (inter-quartile range; 29-55 weeks) weeks and where acceptance and publication dates are known (n=55) median processing time was 18 weeks (inter-quartile range; 13-27 weeks). The difference gives some indication of the time taken in peer review.

The 3 journal-published and 6 Cochrane reviews that reported more than one search date showed shorter lags from last search date to publication than those that did not appear to have updated the search. Eight HTA reviews reported updating their search and 11 did not, but the lags from most recent search to publication were essentially the same. Still, there was a significant overall effect by level of search updating (Log Rank (Mantel-Cox), Chi-square 7.253, df=1, p =0.007).

Publication lags were assessed in the main cohort to examine and trends over time. There was an apparent trend towards decreased publication lags over time, with more recent publication dates having shorter publication lags (p =0.12). However, this reflected bias sampling in the sense that the only way for a recent article to be sampled for inclusion in the cohort would be by having a short publication lag. In other words, systematic reviews initiated in, say, 2004, could only end up in the cohort, if they had relatively short delays before

publication. To avoid this bias, we analyzed the relationship between publication date and publication lag using only systematic reviews published prior to January 1, 2003. In this analysis, the trend towards shorter publication lags with more recent reviews disappeared completely, with a much smaller regression coefficient and p-value > 0.8.

Publication Velocity

The patterns of evidence accumulation at the macro level (by clinical area), or at the micro level (within a particular systematic review) could help to identify or predict optimal update intervals. Velocity at the macro level is considered here.

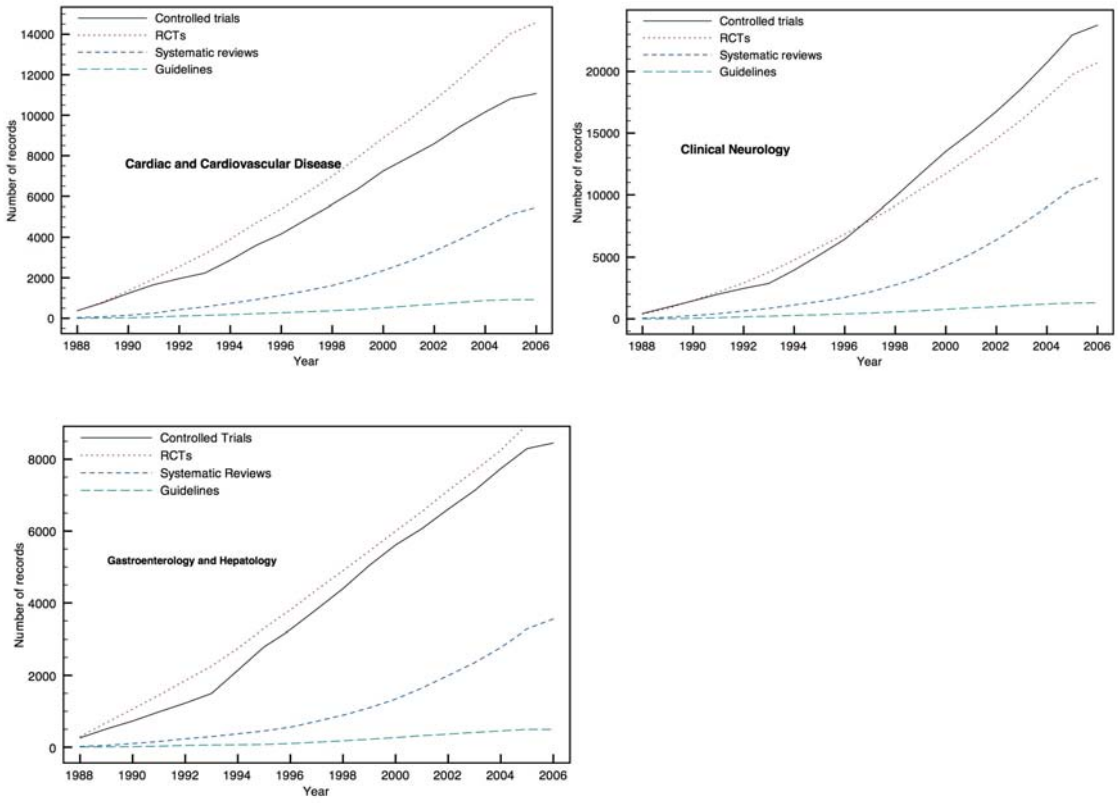


Figure 8. Growth of controlled trials, RCTs, systematic reviews and clinical practice guidelines, 1988-2006.

The three clinical areas with greatest representation in the cohort are cardiac and cardiovascular disease, neurology and gastroenterology (Table 1). Growth of randomized controlled trials, other controlled trials, appear linear in this time frame (Figure 8, Table 6).

Publication doubling times, when calculated under the assumption of linearity, were consistent across clinical content areas and increased markedly as the time from the series start increased (Table 7). All series shown here begin in 1988. For example, approximately 981 new oncology trials were published in 1988. This number doubled in a little over two years (2.2) in 1990 and will take almost 20 years (18.4) for those studies published in 2005.

Table 7. Linear Fit And Rate Of Growth By Clinical Area, Doubling Time In Years From Various Starting Years Assuming Linear Growth.

Clinical area	Pearson R Cumulative RCTs & CCTs	Rate of new RCTs per year	Doubling time from starting year 1990	1995	2000	2005
Oncology	1.00	981	2.2	6.0	11.1	18.4
Clinical Neurology	0.98	1158	1.9	4.4	9.6	18.4
Peripheral Vascular Disease	1.00	1238	2.3	6.4	11.5	18.2
Infectious Diseases	1.00	1227	2.0	6.0	11.4	17.6
Respiratory System	1.00	866	2.0	5.9	11.4	17.8
Cardiac and Cardiovascular Systems	0.99	814	2.1	5.3	10.6	18.3
Psychiatry	0.99	784	1.9	4.7	9.4	18.7
Endocrinology and Metabolism	0.99	581	2.0	3.9	8.8	19.0
Gastroenterology and Hepatology	1.00	515	2.4	8.0	11.5	18.0
Obstetrics and Gynecology	1.00	400	2.1	6.1	11.3	18.1
Urology and Nephrology	1.00	188	2.0	5.6	10.4	18.7
Rheumatology	0.99	157	1.8	4.8	9.6	19.0
Critical Care Medicine	0.99	61	1.3	3.4	10.0	17.2

Policies and Practices of Agencies or Organizations that Fund or Conduct Systematic Reviews

Respondents. Of the 22 Internet surveys sent by email request, 19 organizations responded yielding an overall response rate of 86%, with 17 groups having completed all mandatory questions. Responding organizations were from the U.S., Canada, U.K., and Australia. The majority of respondent organizations identified themselves as producers of systematic reviews (13/19; 68%) with the remainder presenting as both funder and producer combined (4/19; 21%), or exclusively as funders (2/19; 11%). Of those groups surveyed, all indicated they were not-for-profit, and were predominantly academic institutions (9/18; 50%) or national government agencies (5/18; 28%). Government research or infrastructure grants accounted for the majority of funding as reported by groups (16/18; 89%) followed by non-profit academic or non-governmental organization funding (8/18; 44%), internal funding (6/18; 33%) and industry or private sector funding (6/18; 33%).

Main findings. The majority of organizations indicated they produced systematic reviews for the collective goal of both knowledge and decision support (74%; 14/19), while 21% (4/19) reported producing reviews for decision-support. A large portion of respondents (15/19; 79%) view the importance of updating systematic reviews as high to very high. In spite of this however, most organizations do not have a policy in place for updating (13/19; 68%). Nevertheless, of these groups with no formal update processes, 54% (7/13) indicated establishing a policy was of importance. Of those organizations that reportedly update, 68% (13/19) indicate they do so irregularly. Approximately two thirds (13/19; 68%) of respondents reported that at least 20% of the reviews they commission or produce are out of date, and 32% respondents (6/19) reported that at least 50% of their reviews were out of date. When looking at issues of accountability, respondents specified that funder(s) of the original review (5/19; 26%), authors of the original review (5/19; 26%), and policymakers utilizing the evidence (3/19; 16%) were most responsible for ensuring systematic reviews are updated.

The use of formal methods to determine the need to update a systematic review was reported by 32% (6/19) of groups surveyed, 32% (6/19) reported the use of informal methods while an additional 37% (7/19) reportedly use no methods. When looking more in depth at updating strategies and practices, approximately half of organizations do not engage in regular literature searches to identify new evidence (10/19; 53%). However, of those groups that search periodically, searching frequencies were quite variable, with one group reporting monthly searching; two groups reporting every 12 months; one group indicating every two years; and one group stipulating that searching was dependent upon the stability of the evidence base and the relevance of the topic to their audience. The two most frequently reported strategies used (sometimes, often, or always) to monitor the emergence of new evidence were contacting experts in the field (14/18; 78%) and conducting general literature searches including electronic and hand searches (11/19; 58%). Additional surveillance strategies are listed below in Table **8**.

Table 8. Monitoring Strategies

	N of Respondents; %
Experts in the field	14/18; 78%
General literature searches	11/19; 58%
Automatic database alerts or surveillance software	9/17; 53%
Systematic reviews surveillance	9/18; 50%
Guideline or health technology assessment surveillance	7/18; 39%
Trial registry surveillance	7/18; 39%
Statistical approaches	2/18; 11%

When examining updating influences, individuals or groups that reportedly impact most (sometimes, often or always) upon an organization's decisionmaking process of whether to fund or conduct an update are as follows: external policymakers (16/19; 84%); the organization itself as the funder of the systematic review (15/19; 79%); experts in the field (13/19; 68%); and authors of the original review (13/19; 68%). Statisticians (1/19; 5%) and information specialists (3/19; 16%) were least likely to impact this decision. We also note that 26% of groups surveyed indicated that patients or consumer groups 'sometimes' influence this decisionmaking process. When assessing specific issues that may factor into determining 'when' to update, a formal request from a policy or healthcare decisionmaker is the most frequently cited factor by the majority of respondents (16/19; 84%) followed by the totality of all new evidence under consideration (13/17; 76%). See Table 9 for additional impact factors.

Table 9. Factors that Impact on Determining "When" to Update

	N of Respondents; %
Formal request from a policy or healthcare decision maker	16/19; 84%
Totality (comprehensiveness) of all new evidence or data including harms & benefits	13/17; 76%
Number of new studies identified	11/17; 65%
Reporting of serious or 'new' serious adverse events	11/17; 65%
Time credibility	10/18; 56%
Need for an internal organizational decision	8/16; 50%
New inclusion criteria (outcomes; interventions; populations; methodological advances/new analysis)	8/17; 47%
Number of participants in new studies	8/17; 47%

Additional updating influences include the notion that updating will have an effect on clinical practice (15/18; 83%), policy (13/18; 72%), organizational credibility of being current (13/18; 72%), current public controversy or interest (12/18; 67%), or cost utility of updating (12/18; 67%)

Data collected indicates that 60% (9/16; 56%) of respondents spend over 3 months of effort per review on activities related to updating systematic reviews, and 36% (6/16) reported expending over 6 months on updating. When looking closer at type of updating involvement, 72% of groups (13/18) report having 'sometimes' or 'often' been involved in doing full updates of all sections of a review. Two-thirds of respondents report 'often' or 'sometimes' having been involved in partial updates involving only certain sections of original reviews, while 61% of the groups (11/18; 59%) report having been involved in conducting an entirely new review upon updating. Only 1 of 18 respondents (6%) reported ever having discussed the need for a future update in the text of a systematic review. One third (5/17; 29%) of groups have withdrawn at least one systematic review from circulation after assessing the review as out of date. Approximately, 78% of organizations (14/18) reported they are 'often' or 'sometimes' able to draw on the same people involved in the original review. When asked if they had been involved in updating systematic reviews done by others, 61% (11/18) of respondents indicated they 'seldom' or 'never' done this, six groups 'sometimes' had, and only one group reported 'often' updating reviews done by others.

From the data gathered it would seem that most organizations are seldom or not utilizing current existing methods, such as cumulative meta-analytic approaches, when undertaking updating. The most frequently used approach is the time-based approach implying a pre-set updating frequency (7/18; 39%). (See Table 10.)

Table 10. Methods/Procedures

	Use (often or sometimes)/ N of respondents; %	Use (seldom or never)/ N of respondents; %
Time specific approach	7/18; 39%*	9/18; 50%
Bibliometric database entry-date searching	6/18; 33%	10/18; 56%
Editorial strategy with an algorithm of actions	3/18; 17%	12/18; 67%
Cumulative meta-analysis (or extensions)	3/18; 17%	12/18; 67%
Barrowman's identifying the 'null' diagnostic test	0/18; 0%	15/18; 83%

Identifying recent literature published after the date of the last search but before completion of the final systematic review is quite common among those surveyed with 94% (17/18) of organizations reporting this happens 'sometimes' or 'always'. Organizations also report that this information is usually incorporated as an addendum in the review (11/18; 61%), or as a formal revision to the analysis (9/18; 50%).

Updating Barriers. Several elements of original systematic reviews were identified as moderate to serious barriers when updating as reported by respondents including the perceived need to redo data extraction (11/18; 61%); to change the original screening questions (9/18; 50%); to re-assess study quality (9/18; 50%) and to change the original search strategy (8/18; 44%). Further, respondents identified more broad-spectrum barriers (moderate to serious) to updating including limited funding and resources (17/18; 94%); limited academic credit for updating work (11/18; 61%); and limited publishing formats (9/18; 50%). With knowledge of the aforementioned barriers, it should be noted that 72% (13/18) of organizations reported knowing a systematic review was out of date but were not able to commence updating due to lack of resources (e.g. funding, personnel, time).

Harmonization. By harmonization we mean that different groups involved in the funding, conduct, or reporting of systematic reviews would come together and harmonize on issues of conduct, reporting and policy as it relates to updating systematic reviews. A large portion of respondents (11/19; 58%) indicated they 'somewhat' or 'strongly' support centralizing updating efforts across institutions or agencies that produce systematic reviews (i.e., harmonizing updating efforts). There were several perceived benefits (moderate to major) to participating in international harmonization efforts for updating with the foremost being the use of existing resources more efficiently (15/18; 83%). See Table 8 for a list of additional benefits.

Table 11. Major/Moderate Benefits to Harmonization

	N of respondents; %
Use of existing resources more efficiently	15/18; 83%
Potential to minimize duplication of services	14/18; 78%
Access to new information, ideas, materials or other resources	13/15; 72%
Ability to address issues beyond a single organization's domain	11/15; 61%
Share responsibility across organizations for complex/controversial issues	9/18; 50%

Respondents also indicated several barriers to harmonization, including the possible diversion of an organization's funding resources (15/17; 88%) and insufficient human resources

(14/17; 82%). As well, 76% of those surveyed (13/17) viewed perceived delays in working across organizations and possibly diverting the focus of research mandates within organizations (8/17; 47%) as moderate to serious barriers to collaboration. Obstacles aside, 84% (16/19) of the sample indicated they 'somewhat' to 'strongly' favored the development of a central registry of systematic reviews, which would be similar to efforts within the clinical trials community.

Chapter 4. Discussion

Among 100 systematic reviews, qualitative or quantitative signals for updating occurred for 57% (95% CI: 47%- 67%) of the cohort. Median survival free of a signal for updating was 5.5 years (95% CI: 4.6-7.6). However, in 23 cases, a signal for updating occurred within 2 years, in 15 cases the signal occurred in less than 1 year, and, for 7 reviews, signals for updating had already occurred at the time of publication. Cardiovascular medicine, heterogeneity in the original review, and publication of a new trial larger than the previous largest trial were associated with shorter survival times, while inclusion of greater than 13 studies in the original review was associated with increased survival. However, no feature of the original review significantly predicted a signal for updating occurring within 2 years of publication. Using a search protocol combining PubMed Related Articles and a subject search limited with the optimized clinical query search filter for therapies identified almost all new signaling evidence with median screening burden of 71 new records per review.

Signals for updating occurred frequently and within relatively short timelines. While certain features were associated with shorter survival, prediction of the need to update a particular systematic review within specific time frames of interest (e.g., 2 years, as in the main analysis, or 1 or 3 years as checked in sensitivity analyses) does not appear feasible. It is worth emphasizing that this result is unlikely to change with further research. We tested all readily discernible features of systematic reviews with plausible relationships to the need for updating. We found several factors with statistically significant associations with shorter survival, including two with hazard ratios in the range of 2-3, magnitudes that would certainly be of interest epidemiologically. However, as recently highlighted in a discussion of prognostic tools,[22] associations of this magnitude, despite being of epidemiological interest, generally do not give rise to useful prediction tools. The strength of association required for an epidemiological feature by itself to provide a screening test with useful sensitivity and specificity is orders of magnitude higher (i.e., the factor would need to confer a risk of approximately 200-300 fold). It is extremely unlikely that any features of the original systematic review—alone or in combination—would ever increase the risk of a signal for updating within 2 years to such an extent that these factors could usefully identify reviews in need of greater vigilance (i.e., with acceptable positive and negative predictive values). As such, surveillance of the literature for new evidence holds greater promise than relying on features of the original review to identify reviews likely to need updating within short time periods. A preliminary approach is proposed here (below), but may be refined through additional research.

We also evaluated the extent to which growth in the literature varied across broad areas of inquiry, as defined by clinical specialty (e.g., cardiovascular medicine, infectious diseases, obstetrics and gynecology, psychiatry), in order to determine if different clinical areas warranted greater attention with respect to updating. At this broad level, we found that, while the absolute number of new trials published each year does vary quite widely (from a low of 61 new RCTs per year in critical care medicine to a high of 1238 RCTs per year in peripheral vascular diseases), the doubling time for RCTs was surprisingly constant across fields. Linear growth results in ever-longer doubling times for an evidence base, which may bring stability to reviews where shifts in the direction of research are not a complicating factor, i.e., for reviews that could be updated through cumulative meta-analysis.

This combination of wide variation in absolute numbers of trials but minimal variation in rate of growth suggests that clinical fields probably vary in the number of reviews at risk for requiring an update, but that the risk per review does not differ dramatically as a function of clinical field. Importantly, this analysis of clinical fields involved only the rate of production of new evidence at a very broad level. In the cohort analysis, we did find that the specific field of cardiovascular medicine conferred shorter survival time. This finding may reflect features of the field other than the rate of growth, as the rate of 814 new RCTs per year for cardiovascular medicine fell approximately in the middle of the range of values seen across all 13 clinical areas. Number of publications is but one indicator of the amount of new information available. Its appeal it that it is easily counted. Number of new trials, number of patients (new participant ratio), and number of new events are other potentially relevant units of information for determining the need to update. Although the rate of new trial accrual in cardiovascular disease was ranked six among the clinical areas examined, the studies were large compared to other areas (Figure 2).

Alternatively, the similar rates of growth of the literature at the level of broad clinical areas may mask wide variations in growth rates for specific topics within a field. For instance, the rate of new trials per year in cardiology includes the rate for new RCTs in valvular heart disease, which is very small, as well as the rate for new RCTs related to acute coronary syndromes, which is quite high. In fact, in any of the broad clinical fields, there are likely specific topics that have much greater research activity than the average indicated by the broad field as a whole. Systematic reviews of such rapidly changing topics are likely to be challenging and resource-intensive to maintain.

Practical Implications

Many journals, including those with high impact factors, now routinely publish systematic reviews. However, publishing updates of systematic reviews presents challenges because journals, still largely print in format, only have a certain amount of space to provide for reporting systematic reviews. It is unclear whether they will devote any space for publishing updates. Although this situation might be less problematic for electronic journals, at least one of which is committed to publishing systematic review updates,[23] journal 'real estate' available for updating is largely unknown. Print journals might consider publishing updates as "web-only" material. While this would avoid the problem of limited journal 'real estate,' it would still add to the peer review and editing workload for journals. In the case of AHRQ, the EPC program might want to consider developing its own peer reviewed, open access, indexed journal. Such a move might open an important dissemination venue for publishing systematic review updates.

Since its inauguration in 1997 the EPC program has already produced more than 150 evidence reports. And unlike other systematic reviews, which typically focus on a single question, EPC reports usually contain multiple systematic reviews within a single evidence report. Thus, EPC reports may require updating to an even greater extent than indicated by the present analysis, given the multiple topics addressed in each EPC report. With the development of the Medicare Modernization Act and the subsequent development of the clinical effectiveness reviews, and the renewal of the EPC program for another five years, there is likely to be a growing number of completed evidence reports. As with other systematic reviews, the utility of these evidence reports depends on their remaining up to date. Yet, most reviews are not kept up

to date.[1] While it may be possible to commission each EPC with responsibility for keeping their completed evidence reports up to date, there are obvious drawbacks to such an approach. For example, principal authors of some reviews will not remain at the same centers, making updating more complicated. Competing demands on authors' time as well as insufficient resources allocated to updating represent additional barriers.

Alternatively, the EPC program might commission one of its existing EPCs to centralize and harmonize the updating functions for the entire EPC program. Such an approach would likely be cost efficient and relieve pressures on individual EPCs to focus on identifying new evidence. Moreover, centralizing the updating process, an approach the Cochrane Collaboration is currently experimenting with, would facilitate efforts to study and improve the process of updating. This opportunity is important as the current evidence base to inform how and when to update systematic reviews is limited and new approaches and methodologies need to be developed. Through internal EPC knowledge translation activities such developments could be shared with other EPCs and the wider research and health policy community.

For Users of Systematic Reviews

Users of systematic reviews need to recognize that reviews can become out of date within relatively short time frames. Due to the peer review process, including sequential rejections at different journals in the case of most reviews not published in high impact journals, considerable time may elapse between the date of the last electronic search and the time of publication. Although our sample had an average 'lag' period (between the reported last search and eventual publication date) of a little more than 1 year, this average result reflects considerable variation. In fact, over 50% of the cohort had a publication lag of 1.4 years or greater and 25% had a lag of 2.1 years or greater.

To assess the degree to which a given systematic review remains up to date, readers should examine the most recent search date reported. If the search is over two years old, or if readers cannot ascertain this information from the report, then readers should seriously consider the possibility that the review is out of date. In such cases, readers might consider searching for a new systematic review or new primary studies that address the topic of interest. Another option consists of using secondary publication sources, such as the *ACP Journal Club* or *Clinical Evidence*, to identify recently published reviews or primary studies.

For Producers of Systematic Reviews

The finding that 7% of systematic reviews had signals for updating at the time of publication suggests that authors and publishers need to manage production and dissemination times more efficiently. Our data, drawn from a broad range of systematic reviews, including paper-based journal articles, Cochrane reviews, and health technology assessments, revealed that the median time from last search date to publication was 1.4 years, with a 25th percentile of 0.9 years. However, among reviews that explicitly indicated that an update to the search was performed, the 25th percentile was 28 weeks (0.5 years). Achievement of this benchmark across a greater proportion of systematic reviews would produce important increases in survival time.

It is unclear to what extent publishers of systematic reviews can expedite the peer review and publication process for systematic reviews any more than already attempted for submissions of

all types. However, authors can control two of the factors that contribute to publication lags. First, they can update searches prior to submission to capture any evidence that may have emerged during preparation of the review. Second, authors might consider submitting their work to the journals most likely to accept a given review in order to avoid delays due to multiple iterations of the peer review process. When the process of submission and rejection from other journals has resulted in the passage of 1 year or more since the date of the last search, authors should consider updating the search prior to resubmission.

Proposed Surveillance Search Methodology

Searches constructed to conduct systematic reviews typically have low precision (often less than 5%) because they maximize recall in order to avoid missing any relevant evidence. Signal detection against the noise of very low precision makes standard systematic review searches inefficient for the purpose of identifying the need for updating. Good recall of the most influential new evidence would ideally be balanced with a low screening burden. Among the search strategies tested, the combination of two, Related Articles and the optimized Clinical Query, was sufficient to detect all relevant new studies indexed by the date of the updating signal in all but 4 cases. They achieved this recall with precision of approximately 35%. Assessing the performance of our methodology required that we screen candidates sequentially by indexing date in order to identify the earliest point at which a signal could be detected. Using this approach, we found that we needed to screen an average of 71 records per review to detect a signal for updating. In practice, however, there would be no requirement to screen purely in chronological order. For instance, one might first review citations retrieved from the top 5 general journals and the highest impact specialty journals related to the topic of interest. Triaging by journal source would likely reduce the 'number needed to screen' substantially.

In this retrospective assessment, reviewing newer systematic reviews was a useful adjunct to identifying new trials. This technique would have less utility for real time surveillance as lag of at least a year can be expected between trial publication and publication of a systematic review including that trial. Also, relying on newer systematic reviews means that the horse will already be out of the barn, as one would in effect be detecting the need for updating by identifying that an update had in fact been performed. Nonetheless, including both systematic reviews and RCTs adds little work, as the number of newer systematic reviews that require consideration will generally be small.

Review Method

Our approach consisted of applying explicit quantitative and qualitative criteria to new evidence. To ascertain quantitative criteria, we selected a limited number of major outcomes of interest and updated the meta-analytic estimates using fixed effects models. To ascertain qualitative criteria, we monitored for new evidence that changed the certainty of the previous findings, identified new harms or major caveats to the previous findings, among other clinically relevant changes in evidence. We did not compare this method to other possible approaches, but the approach has substantial face validity, and it proved feasible across the cohort of 100 systematic reviews.

As a starting point, we propose that major outcomes be identified for monitoring. The outcomes selected would be sufficiently central to the purpose of the review that signals for changed evidence would warrant undertaking an update. Surveillance searching in the form of the related article protocol and clinical queries would than be initiated, with a search frequency of twice yearly. Search results would than be screened and the robustness of findings from the original review be assessed in the face of the new evidence.

The method could be refined or tailored to different circumstances, including changing the way in which it is operationalized. Certainly the original team of authors for a given systematic review could carry out the same methodology. Alternatively, an agency that supports systematic reviews could carry out some of the surveillance and screening with a team of reviewers, but would need to involve one or more authors from the original review or others with expertise in the relevant content area.

Review Frequency

Agencies that fund systematic reviews should consider how often it is practical to conduct searches, particularly electronic searches, to identify potential new evidence for inclusion in a systematic review. We found that 4% of reviews had a signal for updating within 1 year of the last search date, and that 11% of reviews had signals for updating within 2 years of the last search date. Assuming that 4% is an acceptable risk and that 11% approaches an unacceptable risk, we suggest performance of the electronic searching process approximately once every six months. This frequency would allow sufficient time that, in the event new evidence is identified, reviewers would be able to carry out an update to their existing review along with publication within another 12 to 18 months, bringing the entire process to 2 years or less.

Central or Distributed Surveillance and Updating

For agencies or organizations that sponsor systematic reviews, the functions of surveillance, the decision to update, and actually performing the update could be central or distributed. A central approach would have an editorial group, a specialized team or administrative authority undertake these activities. A distributed approach would have these activities performed by the team that prepared the systematic review. Central surveillance is feasible, and a central approach may be better able to integrate factors other than new evidence, that may influence the decision

to update. Central surveillance permits a standard methodology that may facilitate priority setting when many reviews have signals for update.

Some updating functions, such as determination that a review is not in need of update, can be done centrally. When a major update is needed, it would most efficiently be performed by the authors of the original review.

Review authors could develop and submit an updating brief as part of the initial review – included and excluded files, explicit identification of the 3 newest and 3 largest included trials, and a summary of any information known to them at the time of submission that is pertinent to updating, such as trial registration numbers for ongoing trials.

Format of Update

Our results suggest that the risk factors we identified were not helpful in predicting signals for updating systematic reviews. Surveillance may prove to be more helpful although, as yet, it is unclear which of the several competing surveillance approaches might be most effective. Regardless of which method(s) will turn out to be most effective it is important to remember that these are only signals for updating. Deciding to 'act upon' an updating signal is likely to vary depending on several factors, such as priority setting by the funders or the interest in the topic. One approach to consider is to develop a decision tree model for each funder and/or producer. The details of such an approach are beyond the scope of the current project.

The flexibility and speed of amending electronic products, and the ability to retract obsolete versions makes that format attractive. Costs for both producers and users of electronic products are less than paper copies. Updating efforts should be restricted to the electronic version. Paper versions may be useful for initial dissemination to an important target audiences, however readers could be informed that only the electronic version is updated and so it should be considered the authoritative version.

Surveillance Costs

Life cycle costing of information products, including systematic review, would include updating and retirement costs as well as the initial costs of production and dissemination. In factoring updating into life cycle planning, not only the cost but also the potential to leverage the initial investment by extending the useful life of the systematic review should be considered. We are unable to isolate costs of various aspects of our assessment from costs associated with the methodological research, such as enhanced record keeping and extended search methods. The time requirements described below are estimates.

One experienced searcher with a Masters in Library and Information Science and technical competencies in record manipulation was able to provide abstracts of new evidence for assessment in one working day per systematic review. Time for subsequent surveillance searches on the same systematic reviews would be considerably less. Additional time is needed for provision of full text of articles appearing relevant, which in this cohort was a median of 7 articles (inter-quartile range; 4-24).

The reviewers who assessed the new evidence for updating signals were trained in medicine and clinical epidemiology. Depending on the complexity of the review, the initial orientation to the review and set up for calculation of quantitative signals required one half to one day of effort.

Time required to review candidate studies was dependent on the volume to be assessed. In this project, candidates spanned the period since the search of the original review and included the result sets of 5 search approaches that were under investigation. The median screening burden of the most effective combination of searches was 71 records per review (inter-quartile range 25 – 106) with a median of 7 obtained and assessed in full text format. Reviewing candidates and integrating relevant ones into the quantitative calculations required from one half to three days, but would be less in real-time surveillance. Validation with current expert opinion sources (which would not necessarily be available during real-time surveillance) and preparation of the case summary for review required several hours. Two to five systematic reviews could be considered in a two-hour team conference. Both preparation of the summary briefing and discussion at the team conference was more time consuming for reviews with signals than those without. With set up considered as a one-time cost, a review and summary might take 0.5 days for a review without a signal, and 1.5 days for one with a signal for updating.

Survey

The survey data collected indicate that most organizations that fund and/or conduct systematic reviews research consider updating as important. Unfortunately, most do not have updating policies in place. However, establishing updating procedures is viewed as noteworthy and something organizations concur should be considered. There is strong agreement that the proposed definition of 'update' is a valid explanation of the process, which over time will help to establish common nomenclature for this emerging methodological area.

The majority of survey respondents support the idea of harmonization of updating efforts across groups. Coordinating updating resources across organizations may facilitate performance of regular searches to identify new literature, something that currently occurs inconsistently at best, and may also foster development and evaluation of formal methods to determine the need for updating. Also, of note, most respondents (84%) favored the development of a central registry, analogous to efforts within the clinical trials community, to coordinate updating activities across agencies and review groups.

Even though a high response rate was achieved, the sample denominator was small, as this was a pilot survey. Therefore, interpretation of results should be made with caution. While these survey data will help to establish a baseline of current updating practices and policies of agencies that sponsor systematic reviews, we plan to undertake a similar survey with a larger sample of organizations engaged in funding or producing systematic reviews.

Limitations of the Review

We conducted the survival analysis using a retrospectively assembled cohort to determine how quickly systematic reviews require updating. The use of a retrospective cohort would be unlikely to bias the results of our analysis, but may have made the method appear more feasible than it would have appeared with a prospective cohort. The retrospective approach allowed us to use newer published systematic reviews, which aided in study identification for a substantial minority of cohort reviews. Relying exclusively on identification of new trials as they came out, as would occur with a prospective approach, would likely require greater effort. The

retrospective approach also broadened the availability of published expert opinion to validate our conclusions. Instead of just editorials accompanying the new trials, we could use information from newer textbooks, practice guidelines, and systematic reviews to guide our assessments of the impact of new evidence on the findings of the original review.

We chose systematic reviews indexed in *ACP Journal Club* as our sampling frame in order to construct a cohort of reviews of above average quality that were also directly relevant to clinical practice. It is possible that choosing *ACP Journal Club* introduced a bias in our cohort, such as preferential inclusion of reviews with positive results. A recent comparison of randomized controlled trials summarized in *ACP Journal Club* with the general population of trials indexed in MEDLINE found that *ACP Journal Club* preferentially abstracts randomized controlled trials that report positive results.[24] *ACP Journal Club* might also preferentially focus on systematic reviews with positive results. While this could introduce some imbalance in the sample, such a bias would be unlikely to undermine our findings, as users of systematic reviews likely act on systematic reviews with positive findings more often than ones with null or negative findings. Nonetheless, understanding differences in updating between null reviews and those with clearly positive findings may represent a useful avenue for future methodological research.

We excluded from our analysis all qualitative reviews, reviews of non-therapeutic topics, individual patient data meta-analyses, and meta-regressions, based on our concerns that rates of change in evidence might differ across these different types of reviews. Restriction of our analysis to systematic reviews of randomized trials of conventional drugs, devices, or procedures that reported meta-analytic results for at least one dichotomous outcome may thus seem to have limited generalizability. However, as shown in Figure 2, excluding the records retrieved by our initial electronic search that were not systematic reviews, 139 (48%) of the first 287 systematic reviews were eligible for inclusion. (39 of the 139 eligible reviews were not included because they addressed topics that overlapped with earlier, eligible reviews, but they were nonetheless eligible for inclusion). Thus, while our cohort may appear highly selected, approximately half of the reviews indexed in *ACP Journal Club* were eligible for inclusion in our cohort, reflecting the fact that quantitative reviews of conventional drug therapies represent a substantial proportion of the systematic reviews directly relevant to clinical practice.

Our use of a structured approach for assessing differences between studies of the same topic without involving panels of experts represents the norm in methodological work of this type, including assessments of discrepancies between systematic reviews and large trials,[25,26] variations in results between studies of different designs,[14,27,28] differences in results presented in abstract form versus subsequent journal articles,[12] and highly cited trials versus other trials addressing the same question.[13] Nonetheless, assessments of the need to update previous systematic reviews would ideally include input by content experts who had evaluated the new evidence. As part of a follow-up project, we are conducting such an exercise with a subset of systematic reviews in this cohort to provide validation for the approach used. The idea of quantitative thresholds for indicating the need to update a previous analysis has face validity, but there is no basis for choosing a generic threshold. Therefore, we could have explored the impact on our results of different choices for these thresholds (e.g., using a threshold of 25% for changes in the magnitude of effect estimates, instead of 50%). Several alternatives were identified in a review of systematic review updating methods[4] including optimal information size[29] and new participant ratio.[5] Signals based on these quantities could be developed for real world updating in place of or in addition to these current criteria. In the case of an individual review, such

approaches, especially optimal information size, would be attractive as they would be less arbitrary. However, operationalizing these methods over a cohort of some 100 reviews would require making a number of imputations that would themselves be somewhat arbitrary (e.g., generic values for event rates in control groups and expected effect sizes).

The idea of qualitative changes in evidence that warrant updating a previous review also has face validity, but, again, the specific signals we chose were somewhat arbitrary. However, the concepts captured by our qualitative signals (substantial changes in certainty, new harm, emergence of superior alternative treatments, important caveats about the patient populations who benefit from treatment, and other such issues) emerged from input from our technical advisory panel, as well a published framework for evaluating the need to update clinical practice guidelines.[30] Others might modify the specific criteria we chose, but we believe the qualitative criteria we used speak for themselves as representing clinically relevant changes in evidence.

Finally, this report presents several novel lines of enquiry concerning updating systematic reviews. We have stated what we believe to be logical implications of our findings, but it will be important to have others groups replicate of research and see whether such replication results in findings similar to ours. This is a rich area for research and as well as independently replicating, we strongly encourage others to extend and refine these lines of research.

Conclusions

In a cohort of high quality systematic reviews directly relevant to clinical practice, signals for updating occurred frequently and within relatively short timelines. A number of features significantly affected survival, but none significantly predicted the need for updating within 2 years.

Methods for identifying reviews in need of updating based on surveillance for new evidence hold more promise. Several of the methods tested were feasible, yielding good recall of relevant new evidence with modest screening burdens.

The majority of organizations engaged in the funding or production of systematic reviews view the importance of updating systematic reviews as high to very high. Despite this recognition, most organizations report having no formal policy in place for updating previous systematic reviews. Slightly less than half of organizations performed periodic literature searches to identify new evidence, but searching frequencies varied widely, from monthly to every two years.

If systematic reviews are to achieve their stated goal of providing the best evidence to inform clinical decision making and healthcare policy, issues related to identifying reviews in need of updating will require much greater attention. In the meantime, publishers of systematic reviews should consider a policy of requiring authors to update searches performed over 12 months prior to submission. And, users of systematic reviews need to recognize that important new evidence can appear within short timelines. When considering the results of a particular systematic review, users should search for more recent reviews or trials to see if any exist and determine if the results are consistent with the previous review.

References and Included Studies

Cohort Meta-Analyses Sorted by Signals

Primary signal for updating

Potentially invalidating qualitative signal (with or without quantitative signal)

1. Alderson P, Schierhout G, Roberts I, Bunn F. Colloids versus crystalloids for fluid resuscitation in critically ill patients. *Cochrane Database Syst Rev*. 2000(2):CD000567.
2. Alejandria MM, Lansang MA, Dans LF, Mantaring JB. Intravenous immunoglobulin for treating sepsis and septic shock. *Cochrane Database Syst Rev*. 2002(1):CD001090.
3. Bucher HC, Guyatt GH, Cook RJ, et al. Effect of calcium supplementation on pregnancy-induced hypertension and preeclampsia: a meta-analysis of randomized controlled trials. *JAMA*. 1996;275(14):1113-7.
4. Coward LJ, Featherstone RL, Brown MM. Safety and efficacy of endovascular treatment of carotid artery stenosis compared with carotid endarterectomy: a Cochrane systematic review of the randomized evidence. *Stroke*. 2005;36(4):905-11.
5. Hemminki E, McPherson K. Impact of postmenopausal hormone therapy on cardiovascular events and cancer: pooled data from clinical trials. *BMJ*. 1997;315(7101):149-53.
6. Kjaergard LL, Krogsgaard K, Gluud C. Interferon alfa with or without ribavirin for chronic hepatitis C: systematic review of randomised trials. *BMJ*. 2001;323(7322):1151-5.
7. Lefering R, Neugebauer EA. Steroid controversy in sepsis and septic shock: a meta-analysis. *Crit Care Med*. 1995;23(7):1294-303.
8. Lord JM, Flight IH, Norman RJ. Metformin in polycystic ovary syndrome: systematic review and meta-analysis. *BMJ*. 2003;327(7421):951-3.

Major signal for updating

9. Gotzsche PC, Johansen HK. Meta-analysis of prophylactic or empirical antifungal treatment versus placebo or no treatment in patients with cancer complicated by neutropenia. *BMJ*. 1997;314(7089):1238-44.
10. Abramson MJ, Puy RM, Weiner JM. Is allergen immunotherapy effective in asthma? A meta-analysis of randomized controlled trials. *Am J Respir Crit Care Med*. 1995;151(4):969-74.
11. Collaborative meta-analysis of randomised trials of antiplatelet therapy for prevention of death, myocardial infarction, and stroke in high risk patients. *BMJ*. 2002;324(7329):71-86.
12. Avezum A, Tsuyuki RT, Pogue J, Yusuf S. Beta-blocker therapy for congestive heart failure: a systemic overview and critical appraisal of the published trials. *Can J Cardiol*. 1998;14(8):1045-53.
13. Benavente O, Moher D, Pham B. Carotid endarterectomy for asymptomatic carotid stenosis: a meta-analysis. *BMJ*. 1998;317(7171):1477-80.
14. Birck R, Krzossok S, Markowetz F, Schnulle P, van der Woude FJ, Braun C. Acetylcysteine for prevention of contrast nephropathy: meta-analysis. *Lancet*. 2003;362(9384):598-603.
15. Bischoff-Ferrari HA, Dawson-Hughes B, Willett WC, et al. Effect of Vitamin D on falls: a meta-analysis. *JAMA*. 2004;291(16):1999-2006.
16. Neal B, MacMahon S, Chapman N. Effects of ACE inhibitors, calcium antagonists, and other blood-pressure-lowering drugs: results of prospectively designed overviews of randomised trials. Blood Pressure Lowering Treatment Trialists' Collaboration. *Lancet*. 2000;356(9246):1955-64.
17. Blumenauer B, Judd M, Cranney A, et al. Etanercept for the treatment of rheumatoid arthritis. *Cochrane Database Syst Rev*. 2003(4):CD004525.
18. Bosch X, Marrugat J. Platelet glycoprotein IIb/IIIa blockers for percutaneous coronary revascularization, and unstable angina and non-ST-segment elevation myocardial infarction. *Cochrane Database Syst Rev*. 2001(4):CD002130.
19. Boucher M, McAuley L, Brown A, Keely E, Skidmore B. Efficacy of rosiglitazone and pioglitazone compared to other anti-diabetic agents: systematic review and budget impact analysis. . *Ottawa: Canadian Coordinating*

Office for Health Technology Assessment. Technology report no 29. . 2002.
20. Brown DL, Fann CS, Chang CJ. Meta-analysis of effectiveness and safety of abciximab versus eptifibatide or tirofiban in percutaneous coronary intervention. *Am J Cardiol*. 2001;87(5):537-41.
21. Bucher HC, Griffith LE, Guyatt GH. Effect of HMGcoA reductase inhibitors on stroke. A meta-analysis of randomized, controlled trials. *Ann Intern Med*. 1998;128(2):89-95.
22. Bucher HC, Hengstler P, Schindler C, Guyatt GH. Percutaneous transluminal coronary angioplasty versus medical treatment for non-acute coronary heart disease: meta-analysis of randomised controlled trials. *BMJ*. 2000;321(7253):73-7.
23. Carlberg B, Samuelsson O, Lindholm LH. Atenolol in hypertension: is it a wise choice? *Lancet*. 2004;364(9446):1684-9.
24. Cody J, Daly C, Campbell M, et al. Recombinant human erythropoietin for chronic renal failure anaemia in pre-dialysis patients. *Cochrane Database Syst Rev*. 2005(3):CD003266.
25. Cranney A, Welch V, Adachi JD, et al. Etidronate for treating and preventing postmenopausal osteoporosis. *Cochrane Database Syst Rev*. 2001(4):CD003376.
26. Crouse JR, 3rd, Byington RP, Hoen HM, Furberg CD. Reductase inhibitor monotherapy and stroke prevention. *Arch Intern Med*. 1997;157(12):1305-10.
27. Cullum N, McInnes E, Bell-Syer SE, Legood R. Support surfaces for pressure ulcer prevention. *Cochrane Database Syst Rev*. 2004(3):CD001735.
28. D'Amico R, Pifferi S, Leonetti C, Torri V, Tinazzi A, Liberati A. Effectiveness of antibiotic prophylaxis in critically ill adult patients: systematic review of randomised controlled trials. *BMJ*. 1998;316(7140):1275-85.
29. Ng D, Salvio F, Hicks G. Anti-leukotriene agents compared to inhaled corticosteroids in the management of recurrent and/or chronic asthma in adults and children. *Cochrane Database Syst Rev*. 2004(2):CD002314.
30. Effects of radiotherapy and surgery in early breast cancer. An overview of the randomized trials. Early Breast Cancer Trialists' Collaborative Group. *N Engl J Med*. 1995;333(22):1444-55.
31. Etminan M, Levine MA, Tomlinson G, Rochon PA. Efficacy of angiotensin II receptor antagonists in preventing headache: a systematic overview and meta-analysis. *Am J Med*. 2002;112(8):642-6.
32. Evans BW, Clark WK, Moore DJ, Whorwell PJ. Tegaserod for the treatment of irritable bowel syndrome. *Cochrane Database Syst Rev*. 2004(1):CD003960.
33. Ezekowitz JA, Armstrong PW, McAlister FA. Implantable cardioverter defibrillators in primary and secondary prevention: a systematic review of randomized, controlled trials. *Ann Intern Med*. 2003;138(6):445-52.
34. Fahey T, Stocks N, Thomas T. Quantitative systematic review of randomised controlled trials comparing antibiotic with placebo for acute cough in adults. *BMJ*. 1998;316(7135):906-10.
35. Fowlie PW. Prophylactic indomethacin: systematic review and meta-analysis. *Arch Dis Child Fetal Neonatal Ed*. 1996;74(2):F81-7.
36. Freemantle N, Cleland J, Young P, Mason J, Harrison J. beta Blockade after myocardial infarction: systematic review and meta regression analysis. *BMJ*. 1999;318(7200):1730-7.
37. Gadsby JG, Flowerdew MW. The effectiveness of transcutaneous electrical nerve stimulation (TENS) and acupuncture-like transcutaneous electrical nerve stimulation (ALTENS) in the treatment of patients with chronic low back pain. *Cochrane Database Syst Rev*. 1996.
38. Gotzsche PC, Gjorup I, Bonnen H, Brahe NE, Becker U, Burcharth F. Somatostatin v placebo in bleeding oesophageal varices: randomised trial and meta-analysis. *BMJ*. 1995;310(6993):1495-8.
39. Goulenok C, Bernard B, Cadranel JF, et al. Flumazenil vs. placebo in hepatic encephalopathy in patients with cirrhosis: a meta-analysis. *Aliment Pharmacol Ther*. 2002;16(3):361-72.
40. Goulis J, Leandro G, Burroughs AK. Randomised controlled trials of ursodeoxycholic-acid therapy for primary biliary cirrhosis: a meta-analysis. *Lancet*. 1999;354(9184):1053-60.
41. Griffiths J, Barber VS, Morgan L, Young JD. Systematic review and meta-analysis of studies of the timing of tracheostomy in adult patients undergoing artificial ventilation. *BMJ*. 2005;330(7502):1243.
42. Hart RG, Halperin JL, McBride R, Benavente O, Man-Son-Hing M, Kronmal RA. Aspirin for the primary prevention of stroke and other major vascular events: meta-analysis and hypotheses. *Arch Neurol*. 2000;57(3):326-32.

43. Hood SC, Moher D, Barber GG. Management of intermittent claudication with pentoxifylline: meta-analysis of randomized controlled trials. *Cmaj*. 1996;155(8):1053-9.
44. Hotopf M, Hardy R, Lewis G. Discontinuation rates of SSRIs and tricyclic antidepressants: a meta-analysis and investigation of heterogeneity. *Br J Psychiatry*. 1997;170:120-7.
45. Ioannidis JP, Cappelleri JC, Lau J, et al. Early or deferred zidovudine therapy in HIV-infected patients without an AIDS-defining illness. *Ann Intern Med*. 1995;122(11):856-66.
46. Chen W, Gluud C. Vaccines for preventing hepatitis B in health-care workers. *Cochrane Database Syst Rev*. 2005(4):CD000100.
47. Jefferson T, Deeks JJ, Demicheli V, Rivetti D, Rudin M. Amantadine and rimantadine for preventing and treating influenza A in adults. *Cochrane Database Syst Rev*. 2004(3):CD001169.
48. Jong P, Demers C, McKelvie RS, Liu PP. Angiotensin receptor blockers in heart failure: meta-analysis of randomized controlled trials. *J Am Coll Cardiol*. 2002;39(3):463-70.
49. Keenan SP, Kernerman PD, Cook DJ, Martin CM, McCormack D, Sibbald WJ. Effect of noninvasive positive pressure ventilation on mortality in patients admitted with acute respiratory failure: a meta-analysis. *Crit Care Med*. 1997;25(10):1685-92.
50. Koch A, Bouges S, Ziegler S, Dinkel H, Daures JP, Victor N. Low molecular weight heparin and unfractionated heparin in thrombosis prophylaxis after major surgical intervention: update of previous meta-analyses. *Br J Surg*. 1997;84(6):750-9.
51. Laine L, Schoenfeld P, Fennerty MB. Therapy for Helicobacter pylori in patients with nonulcer dyspepsia. A meta-analysis of randomized, controlled trials. *Ann Intern Med*. 2001;134(5):361-9.
52. Lensing AW, Prins MH, Davidson BL, Hirsh J. Treatment of deep venous thrombosis with low-molecular-weight heparins. A meta-analysis. *Arch Intern Med*. 1995;155(6):601-7.
53. Letelier LM, Udol K, Ena J, Weaver B, Guyatt GH. Effectiveness of amiodarone for conversion of atrial fibrillation to sinus rhythm: a meta-analysis. *Arch Intern Med*. 2003;163(7):777-85.
54. McIntyre PB, Berkey CS, King SM, et al. Dexamethasone as adjunctive therapy in bacterial meningitis. A meta-analysis of randomized clinical trials since 1988. *JAMA*. 1997;278(11):925-31.
55. Michels KB, Yusuf S. Does PTCA in acute myocardial infarction affect mortality and reinfarction rates? A quantitative overview (meta-analysis) of the randomized clinical trials. *Circulation*. 1995;91(2):476-85.

Quantitative signal involving primary outcome but no qualitative signal

56. Anand SS, Yusuf S. Oral anticoagulant therapy in patients with coronary artery disease: a meta-analysis. *JAMA*. 1999;282(21):2058-67.
57. Dalby M, Bouzamondo A, Lechat P, Montalescot G. Transfer for primary angioplasty versus immediate thrombolysis in acute myocardial infarction: a meta-analysis. *Circulation*. 2003;108(15):1809-14.
58. Kong DF, Califf RM, Miller DP, et al. Clinical outcomes of therapeutic agents that block the platelet glycoprotein IIb/IIIa integrin in ischemic heart disease. *Circulation*. 1998;98(25):2829-35.

Any signal

Quantitative signal involving non-primary outcome or any of three pre-defined criteria for changes in the amount or precision of information: increase in total number of trials by ≥ 50%, increase in total number of participants by ≥ 50%, publication of a new trial with sample size ≥ 3 times size of previous largest trial, relative of change ≥ 50% for the width of the 95% confidence interval.

59. Adams NP, Bestall JB, Malouf R, Lasserson TJ, Jones PW Inhaled beclomethasone versus placebo for chronic asthma. *Cochrane Database Syst Rev*. 2005(1):CD002738.
60. Arroll B, Kenealy T. Antibiotics for the common cold and acute purulent rhinitis. *Cochrane Database Syst Rev*. 2005(3):CD000247.
61. Barza M, Ioannidis JP, Cappelleri JC, Lau J. Single or multiple daily doses of aminoglycosides: a meta-analysis. *BMJ*. 1996;312(7027):338-45.
62. Beale RJ, Bryg DJ, Bihari DJ. Immunonutrition in the critically ill: a systematic review of clinical outcome. *Crit Care Med*. 1999;27(12):2799-805.
63. Brocklebank D, Wright J, Cates C. Systematic review of clinical effectiveness of pressurised metered dose inhalers versus other hand held inhaler devices for delivering corticosteroids in asthma. *BMJ*. 2001;323(7318):896-900.

64. Chronicle E, Mulleners W. Anticonvulsant drugs for migraine prophylaxis. *Cochrane Database Syst Rev*. 2004(3):CD003226.
65. de Ferranti SD, Ioannidis JP, Lau J, Anninger WV, Barza M. Are amoxycillin and folate inhibitors as effective as other antibiotics for acute sinusitis? A meta-analysis. *BMJ*. 1998;317(7159):632-7.
66. Del Mar CB, Glasziou PP, Spinks AB. Antibiotics for sore throat. *Cochrane Database Syst Rev*. 2004(2):CD000023.
67. Del Mar C, Glasziou P, Hayem M. Are antibiotics indicated as initial treatment for children with acute otitis media? A meta-analysis. *BMJ*. 1997;314(7093):1526-9.
68. Demicheli V, Rivetti D, Deeks JJ, Jefferson TO. Vaccines for preventing influenza in healthy adults. *Cochrane Database Syst Rev*. 2004(3):CD001269.
69. Ducharme F, Schwartz Z, Hicks G, Kakuma R. Addition of anti-leukotriene agents to inhaled corticosteroids for chronic asthma. *Cochrane Database Syst Rev*. 2004(2):CD003133.
70. Eikelboom JW, Anand SS, Malmberg K, Weitz JI, Ginsberg JS, Yusuf S. Unfractionated heparin and low-molecular-weight heparin in acute coronary syndrome without ST elevation: a meta-analysis. *Lancet*. 2000;355(9219):1936-42.
71. Wahlbeck K, Cheine M, Essali MA. Clozapine versus typical neuroleptic medication for schizophrenia. *Cochrane Database Syst Rev*. 2000(2):CD000059.
72. French RS, Cowan FM, Mansour D, et al. Levonorgestrel-releasing (20 microgram/day) intrauterine systems (Mirena) compared with other methods of reversible contraceptives. *Bjog*. 2000;107(10):1218-25.
73. Gisbert JP, Pajares JM. Systematic review and meta-analysis: is 1-week proton pump inhibitor-based triple therapy sufficient to heal peptic ulcer? *Aliment Pharmacol Ther*. 2005;21(7):795-804.
74. Jull AB, Waters J, Arroll B. Pentoxifylline for treating venous leg ulcers. *Cochrane Database Syst Rev*. 2002(1):CD001733.
75. Kellum JA, J MD. Use of dopamine in acute renal failure: a meta-analysis. *Crit Care Med*. 2001;29(8):1526-31.
76. Laine L, Cook D. Endoscopic ligation compared with sclerotherapy for treatment of esophageal variceal bleeding. A meta-analysis. *Ann Intern Med*. 1995;123(4):280-7.
77. Lengeler C. Insecticide-treated bed nets and curtains for preventing malaria. *Cochrane Database Syst Rev*. 2004(2):CD000363.
78. Macones GA, Berlin M, Berlin JA. Efficacy of oral beta-agonist maintenance therapy in preterm labor: a meta-analysis. *Obstet Gynecol*. 1995;85(2):313-7.

New trials but no signal

79. Bacaltchuk J, Hay P. Antidepressants versus placebo for people with bulimia nervosa. *Cochrane Database Syst Rev*. 2003(4):CD003391.
80. Bell-Syer SE, Hart R, Crawford F, Torgerson DJ, Tyrrell W, Russell I. Oral treatments for fungal infections of the skin of the foot. *Cochrane Database Syst Rev*. 2002(2):CD003584.
81. Bjelakovic G, Nikolova D, Simonetti RG, Gluud C. Antioxidant supplements for prevention of gastrointestinal cancers: a systematic review and meta-analysis. *Lancet*. 2004;364(9441):1219-28.
82. Borgaonkar MR, MacIntosh DG, Fardy JM. A meta-analysis of antimycobacterial therapy for Crohn's disease. *Am J Gastroenterol*. 2000;95(3):725-9.
83. Crowley PA. Antenatal corticosteroid therapy: a meta-analysis of the randomized trials, 1972 to 1994. *Am J Obstet Gynecol*. 1995;173(1):322-35.
84. Cullum N, Nelson EA, Fletcher AW, Sheldon TA. Compression for venous leg ulcers. *Cochrane Database Syst Rev*. 2001(2):CD000265.
85. Eikelboom JW, Quinlan DJ, Douketis JD. Extended-duration prophylaxis against venous thromboembolism after total hip or knee replacement: a meta-analysis of the randomised trials. *Lancet*. 2001;358(9275):9-15.
86. Feigin VL, Rinkel GJ, Algra A, Vermeulen M, van Gijn J. Calcium antagonists in patients with aneurysmal subarachnoid hemorrhage: a systematic review. *Neurology*. 1998;50(4):876-83.
87. Fink HA, Mac Donald R, Rutks IR, Nelson DB, Wilt TJ. Sildenafil for male erectile dysfunction: a systematic review and meta-analysis. *Arch Intern Med*. 2002;162(12):1349-60.
88. Garg R, Yusuf S. Overview of randomized trials of angiotensin-converting enzyme inhibitors on mortality and morbidity in patients with heart failure. Collaborative Group on ACE Inhibitor Trials. *JAMA*. 1995;273(18):1450-6.
89. Heidenreich PA, McDonald KM, Hastie T, et al. Meta-analysis of trials comparing beta-

blockers, calcium antagonists, and nitrates for stable angina. *JAMA*. 1999;281(20):1927-36.
90. Hodson EM, Knight JF, Willis NS, Craig JC. Corticosteroid therapy in nephrotic syndrome: a meta-analysis of randomised controlled trials. *Arch Dis Child*. 2000;83(1):45-51.
91. Holdgate A, Pollock T. Systematic review of the relative efficacy of non-steroidal anti-inflammatory drugs and opioids in the treatment of acute renal colic. *BMJ*. 2004;328(7453):1401.
92. Jackson JL, O'Malley PG, Tomkins G, Balden E, Santoro J, Kroenke K. Treatment of functional gastrointestinal disorders with antidepressant medications: a meta-analysis. *Am J Med*. 2000;108(1):65-72.

No new trials

93. Barr RG, Bourbeau J, Camargo CA, Ram FS. Inhaled tiotropium for stable chronic obstructive pulmonary disease. *Cochrane Database Syst Rev*. 2005(2):CD002876.
94. Edmonds ML, Camargo CA, Brenner BE, Rowe BH. Inhaled steroids in acute asthma following emergency department discharge (Cochrane Review). *In: The Cochrane Library, Issue 1, 2001. Oxford: Update Software.* 2001.
95. Colman I, Brown MD, Innes GD, Grafstein E, Roberts TE, Rowe BH. Parenteral metoclopramide for acute migraine: meta-analysis of randomised controlled trials. *BMJ*. 2004;329(7479):1369-73.
96. Furukawa TA, McGuire H, Barbui C. Meta-analysis of effects and side effects of low dosage tricyclic antidepressants in depression: systematic review. *BMJ*. 2002;325(7371):991.
97. Gourlay SG, Stead LF, Benowitz NL. Clonidine for smoking cessation. *Cochrane Database Syst Rev*. 2004(3):CD000058.
98. Hart RG, Benavente O, McBride R, Pearce LA. Antithrombotic therapy to prevent stroke in patients with atrial fibrillation: a meta-analysis. *Ann Intern Med*. 1999;131(7):492-501.
99. Iorio A, Agnelli G. Low-molecular-weight and unfractionated heparin for prevention of venous thromboembolism in neurosurgery: a meta-analysis. *Arch Intern Med*. 2000;160(15):2327-32.
100. Ko DT, Hebert PR, Coffey CS, Sedrakyan A, Curtis JP, Krumholz HM. Beta-blocker therapy and symptoms of depression, fatigue, and sexual dysfunction. *JAMA*. 2002;288(3):351-7.

References

1. Moher D, Tetzlaff J, Tricco AC et al. Epidemiology and reporting characteristics of systematic reviews. PLoS Med 2007; 4(3):e78.
2. Shojania KG, Bero LA. Taking advantage of the explosion of systematic reviews: an efficient MEDLINE search strategy. Eff Clin Pract 2001; 4(4):157-162.
3. Moher D, Tsertsvadze A. Systematic reviews: when is an update an update? Lancet 2006; 367(9514):881-883.
4. Moher D, Tsertsvadze A, Tricco A et al. When and how to update systematic reviews: a methodological systematic review. Journal of Clinical Epidemiology (in press) 2007;
5. Barrowman NJ, Fang M, Sampson M et al. Identifying null meta-analyses that are ripe for updating. BMC Med Res Methodol 2003; 3(1):13.
6. Sutton AJ, Cooper NJ, Jones DR et al. Evidence-based sample size calculations based upon updated meta-analysis. Stat Med 2006;
7. Shekelle PG, Ortiz E, Rhodes S et al. Validity of the Agency for Healthcare Research and Quality clinical practice guidelines: how quickly do guidelines become outdated? JAMA 2001; 286(12):1461-1467.
8. Higgins, J. P. T. and Green, S.Cochrane Handbook for Systematic Reviews of Interventions 4.2.5. Higgins, J. P. T. and Green, S. Available at: http://www.cochrane.org/resources/handbook/hbook.htm Chichester, UK: John Wiley & Sons, Ltd. 2005.
9. The Cochrane Collaboration. Maintaining your review. The Cochrane Collaboration open learning material Module 19 2002; available at: http://www.cochrane-net.org/openlearning/HTML/ mod19.htm
10. Chalmers I, Enkin M, Keirse MJ. Preparing and updating systematic reviews of randomized controlled trials of health care. Milbank Quarterly 1993; 71(3):411-437.
11. Purpose and Procedure. ACP Journal Club. Available at: http://www.acpjc.org/shared/purpose_and_procedure.htm 2007. Accessed: 2-7-2007.
12. Toma M, McAlister FA, Bialy L et al. Transition from meeting abstract to full-length journal article for randomized controlled trials. JAMA 2006; 295(11):1281-1287.
13. Ioannidis JP. Contradicted and initially stronger effects in highly cited clinical research. JAMA 2005; 294(2):218-228.
14. Ioannidis JP, Haidich AB, Pappa M et al. Comparison of evidence of treatment effects in randomized and nonrandomized studies. JAMA 2001; 286(7):821-830.
15. Survey Monkey. Available at: http://www.surveymonkey.com/ 2006. Accessed: 4-20-2006.
16. Eysenbach G, Wyatt J. Using the Internet for surveys and health research. J Med Internet Res 2002; 4(2):e13.
17. Wyatt JC. When to use web-based surveys. J Am Med Inform Assoc 2000; 7(4):426-429.
18. Dillman DA. Internet and interactive voice response surveys. In: Mail and Internet Surveys: The Tailored Design Method. New York, NY.: John Wiley & Sons Incorporated, 2000.
19. Edwards P, Roberts I, Clarke M et al. Increasing response rates to postal questionnaires: systematic review. BMJ 2002; 324(7347):1183.
20. Leece P, Bhandari M, Sprague S et al. Internet versus mailed questionnaires: a controlled comparison (2). J Med Internet Res 2004; 6(4):e39.
21. Crumley ET, Wiebe N, Cramer K et al. Which resources should be used to identify RCT/CCTs for systematic reviews: a systematic review. BMC Med Res Methodol 2005; 5:24.
22. Ware JH. The limitations of risk factors as prognostic tools. N Engl J Med 2006; 355(25):2615-2617.
23. Many reviews are systematic but some have better systems than others [Editorial]. PLoS Med 2007; 4(3)
24. Carter AO, Griffin GH, Carter TP. A survey identified publication bias in the secondary literature. J Clin Epidemiol 2006; 59(3):241-245.
25. Cappelleri JC, Ioannidis JP, Schmid CH et al. Large trials vs meta-analysis of smaller trials: how do their results compare? JAMA 1996; 276(16):1332-1338.
26. LeLorier J, Gregoire G, Benhaddad A et al. Discrepancies between meta-analyses and subsequent large randomized, controlled trials. N Engl J Med 1997; 337(8):536-542.
27. Benson K, Hartz AJ. A comparison of observational studies and randomized, controlled trials. N Engl J Med 2000; 342(25):1878-1886.

28. Concato J, Shah N, Horwitz RI. Randomized, controlled trials, observational studies, and the hierarchy of research designs. N Engl J Med 2000; 342(25):1887-1892.

29. Pogue JM, Yusuf S. Cumulating evidence from randomized trials: utilizing sequential monitoring boundaries for cumulative meta-analysis. Control Clin Trials 1997; 18(6):580-593.

30. Shekelle P, Eccles MP, Grimshaw JM et al. When should clinical guidelines be updated? BMJ 2001; 323(7305):155-157.

Listing of Excluded Studies

1. ___. Anonymous. ISIS-4: a randomised factorial trial assessing early oral captopril, oral mononitrate, and intravenous magnesium sulphate in 58 050 patients with suspected myocardial infarction. Lancet 1995;345(8951):669-82. Not an SR.

2. Abramson M J, Puy R M, Weiner J M. Allergen immunotherapy for asthma. Cochrane Database of Systematic Reviews 2003;(4). Overlap with earlier SR on same topic.

3. Alberts M J, Hademenos G, Latchaw R E et al. Recommendations for the establishment of primary stroke centers. JAMA 2000;283:3102-9. Not clinical benefit or harm of a specific (class of) drug, device or procedure.

4. Alderson P, Bunn F, Lefebvre C. Human albumin solution for resuscitation and volume expansion in critically ill patients. Cochrane Database of Systematic Reviews 2002;(1). Overlap with earlier SR on same topic.

5. Alejandria M M, Lansang M A, Dans L F et al. Intravenous immunoglobulin for treating sepsis and septic shock. Cochrane Database of Systematic Reviews 2001;(2). Overlap with earlier SR on same topic.

6. Aligne C A, Stoddard J J. Tobacco and children : an economic evaluation of the medical effects of parental smoking. Arch Pediatr Adolesc Med 1998;151:648-53. Review of a non-therapeutic topic.

7. Allen C, Glasziou P, Del Mar C. Bed rest: a potentially harmful treatment needing more careful evaluation. Lancet 1999;354:1229-33. Not clinical benefit or harm of a specific (class of) drug, device or procedure.

8. Alper B S, Lewis P R. Does treatment of acute herpes zoster prevent or shorten postherpetic neuralgia? A systematic review of the literature. J Fam Pract 2000;49:255-64. Focused on ineligible therapy.

9. Alper B S, Lewis P R. Treatment of postherpetic neuralgia: A systematic review of the literature. J Fam Pract 2002;51:121-8. Focused on ineligible therapy.

10. American College, of Physicians. Guidelines for using serum cholesterol, high-density lipoprotein cholesterol, and triglyceride levels as screening tests for preventing coronary heart disease in adults. Ann Intern Med 1996;124(5):515-7. Review of a non-therapeutic topic.

11. Amiodarone Trials, Meta-Analysis Investigators. Effect of prophylactic amiodarone on mortality after acute myocardial infarction and in congestive heart failure: meta-analysis of individual data from 6500 patients in randomised trials. Lancet 1997;350(9089):1417-24. Ineligible MA type.

12. Anderson J W, Johnstone B M, Cook-Newell M E. Meta-analysis of the effects of soy protein intake on serum lipids. N Engl J Med 1995;333(5):276-82. Not clinical benefit or harm of a specific (class of) drug, device or procedure.

13. Annane D, Bellissant E, Bollaert P E. Corticosteroids for severe sepsis and septic shock: a systematic review and meta-analysis. BMJ 2004;329(7464):480-8. Overlap with earlier SR on same topic.

14. Anthonisen N R, Skeans M A, Wise R A. The effects of a smoking cessation intervention on 14.5-year mortality: a randomized clinical trial. Ann Intern Med 2005;142:233-9. Not clinical benefit or harm of a specific (class of) drug, device or procedure.

15. Antonelli M, Conti G, Rocco M. A comparison of noninvasive positive-pressure ventilation and conventional mechanical ventilation in patients with acute respiratory failure. N Engl J Med 1998;339:429-35. Not an SR.

16. Astin J A, Harkness E, Ernst E. The efficacy of "distant healing": a systematic review of randomized trials. Ann Intern Med 2000;132:903-10. Not clinical benefit or harm of a specific (class of) drug, device or procedure.

17. Ausejo M, Saenz A, Pham B. The effectiveness of glucocorticoids in treating croup: meta-analysis. BMJ 1999;319:595-600. Ineligible outcome format.

18. Balas E A, Jaffrey F, Kuperman G J. Electronic communication with patients : evaluation of distance medicine technology. JAMA 1997;278(2):152-9. Not clinical benefit or harm of a specific (class of) drug, device or procedure.

19. Barbui C, Hotopf M, Freemantle N. Selective serotonin reuptake inhibitors versus tricyclic and heterocyclic antidepressants: comparison of drug adherence. Cochrane Database of Systematic Reviews 2000;(4):MA only of non-clinical measures.

20. Barr R G, Rowe B H, Camargo C A. Methylxanthines for exacerbations of chronic obstructive pulmonary disease: meta-analysis of randomised trials. BMJ 2003;327:643-8. Summary statistic was continuous, not WMD.

21. Beilby J J, Silagy C A. Trials of providing costing information to general practitioners: a systematic review. Med J Aust 1997;167(2):89-92. Not clinical benefit or harm of a specific (class of) drug, device or procedure.

22. Bent S, Saint S, Vittinghoff E et al. Antibiotics in acute bronchitis: a meta-analysis. Am J Med 1999;107:62-7. Summary statistic was continuous, not WMD.

23. Croft A M, Garner P. Mefloquine for preventing malaria in non-immune adult travellers.. Cochrane Database of Systematic Reviews 2000;(2):Focused on ineligible therapy.

24. Crowley P. Interventions to prevent, or improve outcome from, labour at or beyond term. Cochrane Database of Systematic Reviews 1997;(2):Focused on ineligible therapy.

25. Crystal E, Connolly S J, Sleik K et al. Interventions on prevention of postoperative atrial fibrillation in patients undergoing heart surgery: A meta-analysis. Circulation 2002;106:75-80. Focused on ineligible therapy.

26. Cullum N, Deeks J, Sheldon T A et al. Beds, mattresses and cushions for pressure sore prevention and treatment. Cochrane Database of Systematic Reviews 1999;Not clinical benefit or harm of a specific (class of) drug, device or procedure.

27. Damoiseaux R A, van Balen F A, Hoes A W et al. Primary care based randomised, double blind trial of amoxicillin versus placebo for acute otitis media in children aged under 2 years. BMJ 2000;320:350-4. Not an SR.

28. Davis D, Thomson O'Brien M A, Freemantle N. Impact of formal continuing medical education : do conferences, workshops, rounds, and other traditional continuing education activities change physician behavior or health care outcomes?. JAMA 1999;282:867-74. Not clinical benefit or harm of a specific (class of) drug, device or procedure.

29. De Bruyn G, Hahn S, Borwick A. Antibiotic treatment for travellers' diarrhoea. Cochrane Database of Systematic Reviews 2000;(3):Other.

30. Dear K, Holden J, Andrews R et al. Vaccines for preventing pneumococcal infection in adults. Cochrane Database of Systematic Reviews 2003;(4). Overlap with earlier SR on same topic.

31. Desai A S, Fang J C, Maisel W H et al. Implantable defibrillators for the prevention of mortality in patients with nonischemic cardiomyopathy: a meta-analysis of randomized controlled trials. JAMA 2004;292:2874-9. Overlap with earlier SR on same topic.

32. Di Blasi Z, Harkness E, Ernst E et al. Influence of context effects on health outcomes: a systematic review. Lancet 2001;357:757-62. Not clinical benefit or harm of a specific (class of) drug, device or procedure.

33. Dick PT, Canadian Task, Force on et al. Periodic health examination, 1996 update: 1 Prenatal screening for and diagnosis of Down syndrome. Can Med Assoc J 1996;154(4):465-79. Review of a non-therapeutic topic.

34. Dickson R, Fullerton D, Eastwood A et al. Preventing and reducing the adverse effects of unintended teenage pregnancies. Effective Health Care 1997;3(1):12. Focused on ineligible therapy.

35. Dimmock P W, Wyatt K M, Jones P W et al. Efficacy of selective serotonin-reuptake inhibitors in premenstrual syndrome: a systematic review. Lancet 2000;356 :1131-6. Ineligible outcome format.

36. Dinnes J, Kleijnen J, Leitner M et al. Cardiac rehabilitation. Health Care 1999;865-71. Not clinical benefit or harm of a specific (class of) drug, device or procedure.

37. Dodds CS. Interventions for treating headlice. Cochrane Database of Systematic Reviews 2000;(2):Focused on ineligible therapy.

38. Dolovich L R, Ginsberg J S, Douketis J D et al. A meta-analysis comparing low-molecular-weight heparins with unfractionated heparin in the treatment of venous thromboembolism : examining some unanswered questions regarding location of treatment, product type, and dosing frequency. Arch Intern Med 2000;160:181-8. Overlap with earlier SR on same topic.

39. Doody R S, Stevens J C, Beck C. Practice parameter: management of dementia (an evidence-based review) : report of the Quality Standards Subcommittee of the American Academy of Neurology. Neurology 2001;56:1154-66. Focused on ineligible therapy.

40. Douketis J D, Feightner J W, Attia J et al. Periodic health examination, 1999 update: 1 Detection,

prevention and treatment of obesity. CMAJ 1999;160:513-25. Focused on ineligible therapy.

41. D'Souza A L, Rajkumar C, Cooke J et al. Probiotics in prevention of antibiotic associated diarrhoea: meta-analysis. BMJ 2002;324:1361-4. Not clinical benefit or harm of a specific (class of) drug, device or procedure.

42. Ducharme F, Schwartz Z, Hicks G et al. Addition of anti-leukotriene agents to inhaled corticosteroids for chronic asthma. Cochrane Database of Systematic Reviews 2004;(2). Overlap with earlier SR on same topic.

43. Duley L, Henderson-Smart D J. Drugs for treatment of very high blood pressure during pregnancy. Cochrane Database of Systematic Reviews 2002;(4):Focused on ineligible therapy.

44. Duley L, Henderson-Smart D, Knight M et al. Antiplatelet drugs for prevention of pre-eclampsia and its consequences: systematic review. BMJ 2001;322:329-33. Overlap with earlier SR on same topic.

45. Dusseldorp E, van Elderen T, Maes S et al. A meta-analysis of psychoeducational programs for coronary heart disease patients. Health Psychol 1999;18:506-19. Not clinical benefit or harm of a specific (class of) drug, device or procedure.

46. Eakin E G, Glasgow R E, Riley K M. Review of primary care-based physical activity intervention studies : effectiveness and implications for practice and future research. J Fam Pract 2000;49:158-68. Focused on ineligible therapy.

47. Early Breast, Cancer Trialists', Collaborative Group. Effects of chemotherapy and hormonal therapy for early breast cancer on recurrence and 15-year survival: an overview of the randomised trials. Lancet 2005;365:1687-717. Focused on ineligible therapy.

48. Ebrahim S, Davey Smith G. Multiple risk factor interventions for primary prevention of coronary heart disease. Cochrane Database of Systematic Reviews 1999;Not clinical benefit or harm of a specific (class of) drug, device or procedure.

49. Edmonds M L, Camargo C A, Pollack C V et al. The effectiveness of inhaled corticosteroids in the emergency department treatment of acute asthma: a meta-analysis. Ann Emerg Med 2002;40:145-54. Overlap with earlier SR on same topic.

50. Edwards A, Unigwe S, Elwyn G et al. Effects of communicating individual risks in screening programmes: Cochrane systematic review. BMJ 2003;327:703-9. Not clinical benefit or harm of a specific (class of) drug, device or procedure.

51. Egger M, Davey Smith G, Stettler C et al. Risk of adverse effects of intensified treatment in insulin-dependent diabetes mellitus: a meta-analysis. Diabet Med 1997;14(11):919-28. Ineligible MA type.

52. Eisenberger M A, Blumenstein B A, Crawford E D. Bilateral orchiectomy with or without flutamide for metastatic prostate cancer. N Engl J Med 1998;339:1036-42. Not an SR.

53. Ernst E. Chelation therapy for peripheral arterial occlusive disease: a systematic review. Circulation 1997;96(3):1031-3. No MA for outcome in abstract.

54. Ernst E. The risk-benefit profile of commonly used herbal therapies: ginkgo, St John's wort, ginseng, echinacea, saw palmetto, and kava. Ann Intern Med 2002;136:42-53. Not clinical benefit or harm of a specific (class of) drug, device or procedure.

55. Feder G, Cryer C, Donovan S et al. Guidelines for the prevention of falls in people over 65. BMJ 2000;321:1007-11. Not clinical benefit or harm of a specific (class of) drug, device or procedure.

56. Fendrick A M, Chernew M E, Hirth R A et al. Alternative management strategies for patients with suspected peptic ulcer disease. Ann Intern Med 1995;123(4):260-8. Not an SR.

57. Ferguson J A, Weinberger M. Case management programs in primary care. J Gen Intern Med 1998;13:123-6. Not clinical benefit or harm of a specific (class of) drug, device or procedure.

58. Fiscella K, Franks P. Cost-effectiveness of the transdermal nicotine patch as an adjunct to physicians' smoking cessation counseling. JAMA 1996;275(16):1247-51. Not an SR.

59. Fleming C, Whitlock E P, Beil T L et al. Screening for abdominal aortic aneurysm: a best-evidence systematic review for the U.S : preventive Services Task Force. Ann Intern Med. 2005;142:203-11. Review of a non-therapeutic topic.

60. Fleming M F, Mundt M P, French M T. Brief physician advice for problem drinkers: long-term efficacy and benefit-cost analysis. Alcohol Clin Exp Res 2002;2636-43. Not an SR.

61. Ford A C, Qume M, Moayyedi P. Helicobacter pylori "test and treat" or endoscopy for managing dyspepsia: an individual patient data meta-analysis. Gastroenterology 2005;128:1838-44. Ineligible MA type.

62. Forster A, Young J, Langhorne P et al. Systematic review of day hospital care for elderly people. BMJ 1999;318:837-41. Focused on ineligible therapy.

63. Fraser W D, Krauss I, Brisson-Carrol G et al. Amniotomy to shorten spontaneous labour. Cochrane Database of Systematic Reviews 1995;(2):Date of original could not be established.

64. Friedrich J O, Adhikari N, Herridge M S et al. Meta-analysis: low-dose dopamine increases urine output but does not prevent renal dysfunction or death. Ann Intern Med 2005;142:510-24. Search date too recent (2005).

65. Gage B F, Cardinalli A B, Albers G W et al. Cost-effectiveness of warfarin and aspirin for prophylaxis of stroke in patients with nonvalvular atrial fibrillation. JAMA 1995;274(23):1839-45. Not an SR.

66. Garbutt J C, West S L, Carey T S et al. Pharmacological treatment of alcohol dependence : a review of the evidence. JAMA 1999;281:1318-25. Focused on ineligible therapy.

67. Gardlund B, for the, Heparin Prophylaxis et al. Randomised, controlled trial of low-dose heparin for prevention of fatal pulmonary embolism in patients with infectious diseases. Lancet 1996;347(9012):1357-61. Not an SR.

68. Gardner A W, Poehlman E T. Exercise rehabilitation programs for the treatment of claudication pain: A meta-analysis. JAMA 1995;274(12):975-80. Not clinical benefit or harm of a specific (class of) drug, device or procedure.

69. Garner P, Gulmezoglu A M. Routine antimalarial drug chemoprophylaxis during pregnancy in endemic malarious areas. Cochrane Database of Systematic Reviews 1996;No MA for outcome in abstract.

70. Gaster B, Hirsch I B. The effects of improved glycemic control on complications in type 2 diabetes. Arch Intern Med 1998;158:134-40. Focused on ineligible therapy.

71. Gelber R D, Cole B F, Goldhirsch A. Adjuvant chemotherapy plus tamoxifen compared with tamoxifen alone for postmenopausal breast cancer: meta-analysis of quality-adjusted survival. Lancet 1996;347(9008):1066-71. Focused on ineligible therapy.

72. Gerstmann D R, Minton S D, Stoddard R A. The Provo Multicenter Early High-frequency Oscillatory Ventilation Trial: improved pulmonary and clinical outcome in respiratory distress syndrome. Pediatrics 1996;98(6 Pt 1):1044-57. Not an SR.

73. Gibson P G, Coughlan J, Abramson M. The effects of self-management education and regular practitioner review in adults with asthma. Cochrane Database of Systematic Reviews 1998;Not clinical benefit or harm of a specific (class of) drug, device or procedure.

74. Gibson P G, Coughlan J, Wilson A J. The effects of limited (information only) patient education programs on the health outcomes of adults with asthma. Cochrane Database of Systematic Reviews 1998;1UpdateNot clinical benefit or harm of a specific (class of) drug, device or procedure.

75. Gibson P G, Henry R L, Coughlan J L. The effect of treatment for gastro-oesophageal reflux on asthma in adults and children. Cochrane Database of Systematic Reviews 1999;Focused on ineligible therapy.

76. Gibson P G, Powell H. Written action plans for asthma: an evidence-based review of the key components. Thorax 2004;5994-9. Not clinical benefit or harm of a specific (class of) drug, device or procedure.

77. Gill D, Hatcher S. A systematic review of the treatment of depression with antidepressant drugs in patients who also have a physical illness. Cochrane Database of Systematic Reviews 1998;Date of original could not be established.

78. Gillespie L D, Gillespie W J, Cumming R et al. Interventions to reduce the incidence of falling in the elderly. Cochrane Database of Systematic Reviews 1998;(1):Focused on ineligible therapy.

79. Glaser R, Herrmann H C, Murphy S A. Benefit of an early invasive management strategy in women with acute coronary syndromes. JAMA 2002;288:3124-9. Not an SR.

80. Glasziou P P, Woodward A J, Mahon C M. Mammographic screening trials for women aged under 50 : a quality assessment and meta-analysis. Med J Aust 1995;162(12):625-9. Review of a non-therapeutic topic.

81. Glennie J, for the, Canadian Coordinating et al. The efficacy of tacrine and the measurement of outcomes in Alzheimer's Disease. CCOHTA Technology Overview: Pharmaceuticals 1997;5No MA for outcome in abstract.

82. Gloaguen V, Cottraux J, Cucherat M et al. A meta-analysis of the effects of cognitive therapy in depressed patients. J Affect Disord 1998;49:59-72. Focused on ineligible therapy.

83. Golden R N, Gaynes B N, Ekstrom R D. The efficacy of light therapy in the treatment of mood

disorders: a review and meta-analysis of the evidence. Am J Psychiatry 2005;162(4):656-62. Not clinical benefit or harm of a specific (class of) drug, device or procedure.

84. Gorelick P B, Sacco R L, Smith D B. Prevention of a first stroke : a review of guidelines and a multidisciplinary consensus statement from the National Stroke Association. JAMA 1999;281:1112-20. Focused on ineligible therapy.

85. Gotzsche P C, Hammarquist C, Burr M. House dust mite control measures in the management of asthma: meta-analysis. BMJ 1998;317:1105-10. Not clinical benefit or harm of a specific (class of) drug, device or procedure.

86. Gotzsche P C, Johansen H K. Meta-analysis of short term low dose prednisolone versus placebo and non-steroidal anti-inflammatory drugs in rheumatoid arthritis. BMJ 1998;316:811-8. Ineligible outcome format.

87. Gotzsche P C, Johansen H K. Routine versus selective antifungal administration for control of fungal infections in patients with cancer. Cochrane Database of Systematic Reviews 2002;(2). Overlap with earlier SR on same topic.

88. Gotzsche P C. Somatostatin analogues for acute bleeding oesophageal varices. Cochrane Database of Systematic Reviews 2002;(1). Overlap with earlier SR on same topic.

89. Grossman E, Messerli F H, Goldbourt U. High blood pressure and diabetes mellitus : are all anti-hypertensive drugs created equal?. Arch Intern Med 2000;160:2447-52. Focused on ineligible therapy.

90. Gueyffier F, Bulpitt C, Boissel J P. Antihypertensive drugs in very old people: a subgroup meta-analysis of randomised controlled trials. Lancet 1999;353 :793-6. Ineligible MA type.

91. Hacke W, Kaste M, Fieschi C et al. Randomised double-blind placebo-controlled trial of thrombolytic therapy with intravenous alteplase in acute ischaemic stroke (ECASS II). Lancet 1998;352:1245-51. Not an SR.

92. Haeusler G, Leitich H, van Trotsenburg M et al. Drug therapy of urinary urge incontinence: a systematic review. Obstet Gynecol 2002;100:1003-16. Focused on ineligible therapy.

93. Handoll H H, Madhok R. Surgical interventions for treating distal radial fractures in adults. Cochrane Database of Systematic Reviews 2001;CD003209. Focused on ineligible therapy.

94. Hankey G J, Sudlow C L, Dunbabin D W. Thienopyridine derivatives (ticlopidine, clopidogrel) versus aspirin for preventing stroke and other serious vascular events in high vascular risk patients. Cochrane Database of Systematic Reviews 1999.Overlap with earlier SR on same topic.

95. Hardy M, Coulter I, Morton S C. S-adenosyl-L-methionine for treatment of depression, osteoarthritis, and liver disease. Evid Rep Technol Assess (Summ) 2002;(Oct.):Not clinical benefit or harm of a specific (class of) drug, device or procedure.

96. Hatala R, Dinh T, Cook D J. Once-daily aminoglycoside dosing in immunocompetent adults: a meta-analysis. Ann Intern Med 1996;124(8):717-25. Overlap with earlier SR on same topic.

97. Hayden M, Pignone M, Phillips C et al. Aspirin for the primary prevention of cardiovascular events: a summary of the evidence for the U.S. Preventive Services Task Force. Ann Intern Med 2002;136:161-72. Overlap with earlier SR on same topic.

98. Hazell P, O'Connell D, Heathcote D et al. Efficacy of tricyclic drugs in treating child and adolescent depression: a meta-analysis. BMJ 1995;310(6984):897-901. Ineligible outcome format.

99. Hebert P R, Gaziano J M, Chan K S et al. Cholesterol lowering with statin drugs, risk of stroke, and total mortality. An overview of randomized trials. JAMA 1997;278:313-21. Overlap with earlier SR on same topic.

100. Herxheimer A, Petrie K J. Melatonin for preventing and treating jet lag. Cochrane Database of Systematic Reviews 2001;(1):Date of original could not be established.

101. Hochberg M C, Tracy J K, Flores R H. "Stepping-up" from methotrexate: a systematic review of randomised placebo controlled trials in patients with rheumatoid arthritis with an incomplete response to methotrexate. Ann Rheum Dis 2001;60 : iii51-iii54. Focused on ineligible therapy.

102. Hodnett E D. Home-like versus conventional institutional settings for birth. Cochrane Database of Systematic Reviews 2000;(2):Not clinical benefit or harm of a specific (class of) drug, device or procedure.

103. Hooper L, Brown T J, Elliott R. The effectiveness of five strategies for the prevention of gastrointestinal toxicity induced by non-steroidal anti-inflammatory drugs: systematic review. BMJ 2004;329948Focused on ineligible therapy.

104. Hotopf M, Lewis G, Normand C. Are SSRIs a cost-effective alternative to tricyclics?. Br J Psychiatry 1996;168(4):404-9. Review of a non-therapeutic topic.

105. Hrobjartsson A, Gotzsche P C. Is the placebo powerless? An analysis of clinical trials comparing placebo with no treatment. N Engl J Med 2001;344:1594-602. Focused on ineligible therapy.

106. Huang E S, Meigs J B, Singer D E. The effect of interventions to prevent cardiovascular disease in patients with type 2 diabetes mellitus. Am J Med 2001;111:633-42. Focused on ineligible therapy.

107. Hughes R A, Wijdicks E F, Barohn R. Practice parameter: Immunotherapy for Guillain-Barre syndrome: Report of the Quality Standards Subcommittee of the American Academy of Neurology. Neurology 2003;61:736-40. Focused on ineligible therapy.

108. Hull R D, Pineo G F, Stein P D. Extended out-of-hospital low-molecular-weight heparin prophylaxis against deep venous thrombosis in patients after elective hip arthroplasty: a systematic review. Ann Intern Med 2001;135:858-69. Overlap with earlier SR on same topic.

109. Hull R D, Raskob G E, Brant R F et al. Low-molecular-weight heparin vs heparin in the treatment of patients with pulmonary embolism. Arch Intern Med 2000;160:229-36. Not an SR.

110. Humphrey L L, Chan B K, Sox H C. Postmenopausal hormone replacement therapy and the primary prevention of cardiovascular disease. Ann Intern Med 2002;137:273-84. Non-RCTs included

111. Humphrey L L, Helfand M, Chan B K et al. Breast cancer screening: a summary of the evidence for the U.S. Preventive Services Task Force. Ann Intern Med 2002;137:347-60. Review of a non-therapeutic topic.

112. Hunt D L, Haynes R B, Hanna S E et al. Effects of computer-based clinical decision support systems on physician performance and patient outcomes: A systematic review. JAMA 1998;280:1339-46. Not clinical benefit or harm of a specific (class of) drug, device or procedure.

113. Huntley A, Ernst E. Herbal medicines for asthma: a systematic review. Thorax 2000;55:925-9. Not clinical benefit or harm of a specific (class of) drug, device or procedure.

114. Indredavik B, Bakke F, Slordahl S A et al. Stroke unit treatment improves long-term quality of life: A randomized controlled trial. Stroke 1998;29:895-9. Not an SR.

115. Indredavik B, Slordahl S A, Bakke F et al. Stroke unit treatment. Long-term effects. Stroke 1997;28:1861-6. Not an SR.

116. Ioannidis J P, Cappelleri J C, Skolnik P R et al. A meta-analysis of the relative efficacy and toxicity of Pneumocystis carinii prophylactic regimens. Arch Intern Med 1996;156(2):177-88. Focused on ineligible therapy.

117. Ioannidis J P, Lau J. Evidence on interventions to reduce medical errors. An overview and recommendations for future research. J Gen Intern Med 2001;16:325-34. Not clinical benefit or harm of a specific (class of) drug, device or procedure.

118. Jadad A R, Boyle M, Cunningham C et al. Treatment of attention-deficit/hyperactivity disorder. Evid Rep Technol Assess (Summ) 1999;1-341. Focused on ineligible therapy.

119. Jafar T H, Schmid C H, Landa M et al. Angiotensin-converting enzyme inhibitors and progression of nondiabetic renal disease. A meta-analysis of patient-level data. Ann Intern Med 2001;13:573-87. Ineligible MA type.

120. Jailwala J, Imperiale T F, Kroenke K. Pharmacologic treatment of the irritable bowel syndrome: a systematic review of randomized, controlled trials. Ann Intern Med 2000;133:136-47. Focused on ineligible therapy.

121. Jewell D, Young G. Interventions for nausea and vomiting in early pregnancy. Cochrane Database of Systematic Reviews 2002;(1):Focused on ineligible therapy.

122. Johnson E S, Lanes S F, Wentworth C E. A metaregression analysis of the dose-response effect of aspirin on stroke. Arch Intern Med 1999;159:1248-53. Ineligible MA type.

123. Jones A, Fay J K, Burr M. Inhaled corticosteroid effects on bone metabolism in asthma and mild chronic obstructive pulmonary disease. Cochrane Database of Systematic Reviews 2002;(1):Summary statistic was continuous, not WMD.

124. Jones G, Halbert J, Crotty M. The effect of treatment on radiological progression in rheumatoid arthritis: a systematic review of randomized placebo-controlled trials. Rheumatology 2003;42(6):13. Focused on ineligible therapy.

125. Julian D G, Camm A J, Frangin G et al. Randomised trial of effect of amiodarone on

mortality in patients with left-ventricular dysfunction after recent myocardial infarction: EMIAT. Lancet 1997;349:667-74. Not an SR.

126. Jung A C, Staiger T, Sullivan M. The efficacy of selective serotonin reuptake inhibitors for the management of chronic pain. J Gen Intern Med 1997;12:384-9. No MA for outcome in abstract.

127. Kahan M, Wilson L, Becker L. Effectiveness of physician-based interventions with problem drinkers: a review. CMAJ 1995;152(6):851-9. Not clinical benefit or harm of a specific (class of) drug, device or procedure.

128. Katlama C, Ingrand D, Loveday C et al. Safety and efficacy of lamivudine-zidovudine combination therapy in antiretroviral-naive patients. A randomized controlled comparison with zidovudine monotherapy. JAMA 1996;276(2):118-25. Not an SR.

129. Kattlove H, Liberati A, Keeler E et al. Benefits and costs of screening and treatment for early breast cancer : development of a basic benefit package. JAMA 1995;273(2):142-8. Review of a non-therapeutic topic.

130. Kaushal R, Shojania K G, Bates D W. Effects of computerized physician order entry and clinical decision support systems on medication safety: a systematic review. Arch Intern Med 2003;163:1409-16. Not clinical benefit or harm of a specific (class of) drug, device or procedure.

131. Kearon C, Hirsh J. Starting prophylaxis for venous thromboembolism postoperatively. Arch Intern Med 1995;155(4):366-372. Focused on ineligible therapy.

132. Keeley E C, Boura J A, Grines C L. Primary angioplasty versus intravenous thrombolytic therapy for acute myocardial infarction: a quantitative review of 23 randomised trials. Lancet 2003;361:13-20. Overlap with earlier SR on same topic.

133. Keenan S P, Sinuff T, Cook D J et al. Does noninvasive positive pressure ventilation improve outcome in acute hypoxemic respiratory failure? A systematic review. Crit Care Med 2004;32:2516-23. Overlap with earlier SR on same topic.

134. Kenyon S L, Taylor D J, Tarnow-Mordi W et al. Broad-spectrum antibiotics for preterm, prelabour rupture of fetal membranes: the ORACLE I randomised trial. Lancet 2001;357:979-88. Not an SR.

135. Kenyon S L, Taylor D J, Tarnow-Mordi W et al. Broad-spectrum antibiotics for spontaneous preterm labour: the ORACLE II randomised trial. Lancet 2001;357:989-94. Not an SR.

136. Kerlikowske K, Grady D, Rubin S M et al. Efficacy of screening mammography: A meta-analysis. JAMA 1995;273(2):149-54. Review of a non-therapeutic topic.

137. Kleijnen J, Mackerras D. Vitamin E for the treatment of intermittent claudication. Cochrane Database of Systematic Reviews 1998;(1). UpdateUnneeded Cochrane review.

138. Ko D T, Hebert P R, Coffey C S. Adverse effects of [beta]-blocker therapy for patients with heart failure: a quantitative overview of randomized trials. Arch Intern Med 2004;164:1389-94. Overlap with earlier SR on same topic.

139. Kolbach D N, Sandbrink M W, Hamulyak K et al. Non-pharmaceutical measures for prevention of post-thrombotic syndrome. Cochrane Database of Systematic Reviews 2004;(1). Unneeded Cochrane review.

140. Kronenberg F, Fugh-Berman A. Complementary and alternative medicine for menopausal symptoms: a review of randomized, controlled trials. Ann Intern Med 2002;137:805-13. Not clinical benefit or harm of a specific (class of) drug, device or procedure.

141. Krueger W A, Lenhart F P, Neeser G. Influence of combined intravenous and topical antibiotic prophylaxis on the incidence of infections, organ dysfunctions, and mortality in critically ill surgical patients: a prospective, stratified, randomized, double-blind, placebo-controlled clinical trial. Am J Respir Crit Care Med 2002;166:1029-37. Not an SR.

142. Krystal J H, Cramer J A, Krol W F et al. Naltrexone in the treatment of alcohol dependence. N Engl J Med 2001;345:1734-9. Not an SR.

143. Kwan J, Sandercock P. In-hospital care pathways for stroke. Cochrane Database of Systematic Reviews 2002;(2). Focused on ineligible therapy.

144. Lancaster T, Stead L, Silagy C et al. Effectiveness of interventions to help people stop smoking: findings from the Cochrane Library. BMJ 2000;321:355-8. Focused on ineligible therapy.

145. Langhorne P, Taylor G, Murray G. Early supported discharge services for stroke patients: a meta-analysis of individual patients' data. Lancet 2005;365:501-6. Focused on ineligible therapy.

146. Law M, Tang J L. An analysis of the effectiveness of interventions intended to help people stop smoking. Arch Intern Med 1995;155(18):1933-41. Focused on ineligible therapy.

147. Lechat P, Packer M, Chalon S. Clinical effects of [beta]-adrenergic blockade in chronic heart failure: A meta-analysis of double-blind, placebo-controlled, randomized trials. Circulation 1998;98:1184-91. Overlap with earlier SR on same topic.

148. Lee A, Cooper M G, Craig J C et al. Effects of nonsteroidal anti-inflammatory drugs on post-operative renal function in adults. Cochrane Database of Systematic Reviews 2001;(2):Unneeded Cochrane review.

149. Lees A J, Katzenschlager R, Head J et al. Ten-year follow-up of three different initial treatments in de-novo PD: A randomized trial. Neurology 2001;57:1687-94. Not an SR.

150. Leonardi-Bee J, Bath P M, Bousser M G. Dipyridamole for preventing recurrent ischemic stroke and other vascular events: a meta-analysis of individual patient data from randomized controlled trials. Stroke 2005;36:162-8. Ineligible MA type.

151. Levine C B, Fahrbach K R, Siderowf A D. Diagnosis and treatment of Parkinson's disease: a systematic review of the literature. Evid Rep Technol Assess (Summ) 2003;571-4. Summary statistic was continuous, not WMD.

152. Levine R J, Hauth J C, Curet L B. Trial of calcium to prevent pre-eclampsia. N Engl J Med 1997;337:69-76. Not an SR.

153. Lewin S A, Skea Z C, Entwistle V et al. Interventions for providers to promote a patient-centred approach in clinical consultations. Cochrane Database of Systematic Reviews 2001;(4):Not clinical benefit or harm of a specific (class of) drug, device or procedure.

154. Li Z, Maglione M, Tu W. Meta-analysis: pharmacologic treatment of obesity. Ann Intern Med 2005;142:532-46. Focused on ineligible therapy.

155. Linden W, Stossel C, Maurice J. Psychosocial interventions for patients with coronary artery disease: A meta-analysis. Arch Intern Med 1996;156(7):745-52. Focused on ineligible therapy.

156. Liver Infusion, Meta-analysis Group. Portal vein chemotherapy for colorectal cancer: a meta-analysis of 4000 patients in 10 studies. J Natl Cancer Inst 1997;89:497-505. Ineligible MA type.

157. Low D E, Desrosiers M, McSherry J. A practical guide for the diagnosis and treatment of acute sinusitis. CMAJ 1997;156(Suppl. 6):1S-14S. Focused on ineligible therapy.

158. ls-Nielsen B, Gluud L L, Gluud C. Nonabsorbable disaccharides for hepatic encephalopathy. Cochrane Database of Systematic Reviews 2004;(2). Unneeded Cochrane review.

159. Luria J W, Gonzalez-del-Rey J A, DiGiulio G A. Effectiveness of oral or nebulized dexamethasone for children with mild croup. Arch Pediatr Adolesc Med 2001;155:1340-5. Not an SR.

160. Magee L A, Ornstein M P, von Dadelszen P. Management of hypertension in pregnancy. BMJ 1999;318:1332-6. Focused on ineligible therapy.

161. Maggard M A, Shugarman L R, Suttorp M. Meta-analysis: surgical treatment of obesity. Ann Intern Med 2005;142:547-59. Focused on ineligible therapy.

162. Manser R L, Irving L B, Byrnes G. Screening for lung cancer: a systematic review and meta-analysis of controlled trials. Thorax 2003;58:784-9. Review of a non-therapeutic topic.

163. Man-Son-Hing M, Wells G. Meta-analysis of efficacy of quinine for treatment of nocturnal leg cramps in elderly people. BMJ 1995;310(6971):13-7. Ineligible MA type.

164. Marshall J K, Blackhouse G, Goeree R. Infliximab for the treatment of Crohn's disease: a systematic review and cost-utility analysis. Evid Rep Technol Assess (Summ) 2002;No MA for outcome in abstract.

165. Martin-Hirsch P, Lilford R, Jarvis G et al. Efficacy of cervical-smear collection devices: a systematic review and meta-analysis. Lancet 1999;354:1763-70. Review of a non-therapeutic topic.

166. Mason J, O'Keeffe C, McIntosh A. A systematic review of foot ulcer in patients with type 2 diabetes mellitus : I: Prevention; II: Treatment. Diabet Med 1999;16:801-12. Focused on ineligible therapy.

167. Mayo-Smith M F, Beecher L H, Fischer T L. Management of alcohol withdrawal delirium: an evidence-based practice guideline. Arch Intern Med 2004;164:1405-12. Focused on ineligible therapy.

168. Mayo-Smith M F, for the, American Society et al. Pharmacological management of alcohol withdrawal: A meta-analysis and evidence-based practice guideline. JAMA 1997;278:144-51. Focused on ineligible therapy.

169. McAlindon T E, LaValley M P, Gulin J P et al. Glucosamine and chondroitin for treatment of osteoarthritis: a systematic quality assessment and meta-analysis. JAMA 2000;283 :1469-75. Ineligible outcome format.

170. McAlister F A, Ezekowitz J A, Wiebe N. Systematic review: cardiac resynchronization in patients with symptomatic heart failure. Ann Intern Med 2004;141:381-90. Overlap with earlier SR on same topic.

171. McAlister F A, Lawson F M, Teo K K et al. Randomised trials of secondary prevention programmes in coronary heart disease: systematic review. BMJ 2001;323:957-62. Focused on ineligible therapy.

172. McCartney P, Macdowall W, Thorogood M. A randomised controlled trial of feedback to general practitioners of their prophylactic aspirin prescribing. BMJ 1997;315:35-6. Not an SR.

173. McCrory D C, Brown C D. Anti-cholinergic bronchodilators versus beta2-sympathomimetic agents for acute exacerbations of chronic obstructive pulmonary disease. Cochrane Database of Systematic Reviews 2002;(4):Unneeded Cochrane review.

174. McCusker J, Cole M, Keller E et al. Effectiveness of treatments of depression in older ambulatory patients. Arch Intern Med 1998;158:705-12. Focused on ineligible therapy.

175. McDermott M M, Schmitt B, Wallner E. Impact of medication nonadherence on coronary heart disease outcomes: a critical review. Arch Intern Med 1997;157:1921-9. Review of a non-therapeutic topic.

176. McDonald H P, Garg A X, Haynes R B. Interventions to enhance patient adherence to medication prescriptions: scientific review. JAMA 2002;288:2868-79. Review of a non-therapeutic topic.

177. McLeod RS, members of, the Canadian et al. Screening strategies for colorectal cancer: a systematic review of the evidence. Can J Gastroenterol 2001;15:647-60. Review of a non-therapeutic topic.

178. McQuay H J. Pre-emptive analgesia: a systematic review of clinical studies. Ann Med 1995;27(2):249-56. Focused on ineligible therapy.

179. Mehta S R, Yusuf S, Diaz R. Effect of glucose-insulin-potassium infusion on mortality in patients with acute ST-segment elevation myocardial infarction: the CREATE-ECLA randomized controlled trial. JAMA 2005;293:437-46. Not an SR.

180. MERIT-HF Study, Group. Effect of metoprolol CR/XL in chronic heart failure: Metoprolol CR/XL Randomised Intervention Trial in Congestive Heart Failure (MERIT-HF). Lancet 1999;353:2001-7. Not an SR.

181. Merkel C, Marin R, Enzo E et al. Randomised trial of nadolol alone or with isosorbide mononitrate for primary prophylaxis of variceal bleeding in cirrhosis. Lancet 1996;348(9043):1677-81. Not an SR.

182. Mermel L A. Prevention of intravascular catheter-related infections. Ann Intern Med 2000;132:391-402. Review of a non-therapeutic topic.

183. Messerli F H, Grossman E, Goldbourt U. Are [beta]-blockers efficacious as first-line therapy for hypertension in the elderly? A systematic review. JAMA 1998;279:1903-7. Unneeded Cochrane review.

184. Miller A B, To T, Baines C J et al. The Canadian National Breast Screening Study-1: breast cancer mortality after 11 to 16 years of follow up: A randomized screening trial of mammography in women age 40 to 49 years. Ann Intern Med 2002;137:305-12. Not an SR.

185. Miller J, Chan B K, Nelson H D. Postmenopausal estrogen replacement and risk for venous thromboembolism: a systematic review and meta-analysis for the U.S Preventive Services Task Force. Ann Intern Med 2002;136:680-90. Non-RCTs included

186. Miller M R, McNamara R L, Segal J B. Efficacy of agents for pharmacologic conversion of atrial fibrillation and subsequent maintenance of sinus rhythm: A meta-analysis of clinical trials. J Fam Pract 2000;49:1033-46. Focused on ineligible therapy.

187. Miller R G, Mitchell J D, Moore D H. Riluzole for amyotrophic lateral sclerosis (ALS)/motor neuron disease (MND). Cochrane Database of Systematic Reviews 1999;Date of original could not be established.

188. Milne S, Welch V, Brosseau L. Transcutaneous electrical nerve stimulation (TENS) for chronic low back pain. Cochrane Database of Systematic Reviews 2001;(2). Overlap with earlier SR on same topic.

189. Minneci P C, Deans K J, Banks S M et al. Meta-analysis: the effect of steroids on survival and shock during sepsis depends on the dose. Ann Intern Med

2004;141:47-56. Overlap with earlier SR on same topic.

190. Miyasaki J M, Martin W, Suchowersky O et al. Practice parameter: initiation of treatment for Parkinson's disease: an evidence-based review. Report of the Quality Standards Subcommittee of the American Academy of Neurology. Neurology 2002;58:11-7. No MA for outcome in abstract.

191. Moayyedi P, Delaney B C, Vakil N et al. The efficacy of proton pump inhibitors in nonulcer dyspepsia: a systematic review and economic analysis. Gastroenterology 2004;127:1329-37. Overlap with earlier SR on same topic.

192. Moore R A. Livial: a review of clinical studies. Br J Obstet Gynaecol 1999;106(Suppl):1-21. Focused on ineligible therapy.

193. Morrison L J, Verbeek P R, McDonald A C et al. Mortality and prehospital thrombolysis for acute myocardial infarction: A meta-analysis. JAMA 2000;283:2686-92. Not clinical benefit or harm of a specific (class of) drug, device or procedure.

194. Morrison R S, Chassin M R, Siu A L. The medical consultant's role in caring for patients with hip fracture. Ann Intern Med 1998;128:1010-20. Focused on ineligible therapy.

195. Moyer A, Finney J W, Swearingen C E et al. Brief interventions for alcohol problems: a meta-analytic review of controlled investigations in treatment-seeking and non-treatment-seeking populations. Addiction 2002;97:279-92. Focused on ineligible therapy.

196. Mukherjee D, Nissen S E, Topol E J. Risk of cardiovascular events associated with selective COX-2 inhibitors. JAMA 2001;286:954-9. No MA for outcome in abstract.

197. Mullen P D, Simons-Morton D G, Ram'rez G. A meta-analysis of trials evaluating patient education and counseling for three groups of preventive health behaviors. Patient Educ Couns 1997;32:157-73. Not clinical benefit or harm of a specific (class of) drug, device or procedure.

198. NHS Centre, for Reviews, Dissemination Nuffield et al. Preventing falls and subsequent injury in older people. Eff Health Care 1996;2:1-16. Focused on ineligible therapy.

199. Nista E C, Candelli M, Cremonini F. Levofloxacin-based triple therapy vs. quadruple therapy in second-line Helicobacter pylori treatment: a randomized trial. Aliment Pharmacol Ther 2003;18:627-33. Not an SR.

200. Norris S L, Engelgau M M, Venkat Narayan K M. Effectiveness of self-management training in type 2 diabetes: A systematic review of randomized controlled trials. Diabetes Care 2001;24:561-87. Not clinical benefit or harm of a specific (class of) drug, device or procedure.

201. Norris S L, Zhang X, Avenell A. Efficacy of pharmacotherapy for weight loss in adults with type 2 diabetes mellitus: a meta-analysis. Arch Intern Med 2004;164:1395-404. Focused on ineligible therapy.

202. O'Brien B, Goeree R, Mohamed A H et al. Cost-effectiveness of Helicobacter pylori eradication for the long-term management of duodenal ulcer in Canada. Arch Intern Med 1995;155(18):1958-64. Review of a non-therapeutic topic.

203. O'Connor A M, Rostom A, Fiset V. Decision aids for patients facing health treatment or screening decisions: systematic review. BMJ 1999;319:731-4. Not clinical benefit or harm of a specific (class of) drug, device or procedure.

204. Ofman J J, Rabeneck L. The effectiveness of endoscopy in the management of dyspepsia: a qualitative systematic review. Am J Med 1999;106:335-46. Review of a non-therapeutic topic.

205. Oldman A D, Smith L A, McQuay H J et al. Pharmacological treatments for acute migraine: quantitative systematic review. Pain 2002;97:247-57. Focused on ineligible therapy.

206. O'Mallcy P G, Balden E, Tomkins G. Treatment of fibromyalgia with antidepressants: A meta-analysis. J Gen Intern Med 2000;15:659-66. Ineligible outcome format.

207. O'Malley P G, Jackson J L, Santoro J. Antidepressant therapy for unexplained symptoms and symptom syndromes. J Fam Pract 1999;48:980-90. Ineligible outcome format.

208. O'Meara S M, Cullum N A, Majid M et al. Systematic review of antimicrobial agents used for chronic wounds. Br J Surg 2001;884-21. Focused on ineligible therapy.

209. Ostermann M E, Keenan S P, Seiferling R A et al. Sedation in the intensive care unit: A systematic review. JAMA 2000;283:1451-9. Focused on ineligible therapy.

210. Overgaard M, Hansen P S, Overgaard J et al. Postoperative radiotherapy in high-risk premenopausal women with breast cancer who receive adjuvant chemotherapy. N Engl J Med 1997;337:949-55. Not an SR.

211. Padwal R, Li S K, Lau D C. Long-term pharmacotherapy for obesity and overweight. Cochrane Database of Systematic Reviews 2004;(3):Focused on ineligible therapy.

212. Padwal R, Majumdar S R, Johnson J A et al. A systematic review of drug therapy to delay or prevent type 2 diabetes. Diabetes Care 2005;28:736-44. Focused on ineligible therapy.

213. Parameswaran K, Belda J, Rowe B H. Addition of intravenous aminophylline to beta2-agonists in adults with acute asthma. Cochrane Database of Systematic Reviews 2000;(4):Unneeded Cochrane review.

214. Paramothayan N S, Jones P W. Corticosteroids for pulmonary sarcoidosis. Cochrane Database of Systematic Reviews 2000;(1). Unneeded Cochrane review.

215. Parker M J, Handoll H H, Robinson C M. Gamma nail versus sliding hip screw for the treatment of extracapsular femoral fractures. Cochrane Database of Systematic Reviews 1996;(1):Date of original could not be established.

216. Parkes J, Shepperd S. Discharge planning from hospital to home. Cochrane Database of Systematic Reviews 2000;(4):Review of a non-therapeutic topic.

217. Peter J V, Moran J L, Phillips-Hughes J et al. Noninvasive ventilation in acute respiratory failure-a meta-analysis update. Crit Care Med 2002;30:555-62. Overlap with earlier SR on same topic.

218. Piccinelli M, Pini S, Bellantuono C et al. Efficacy of drug treatment in obsessive-compulsive disorder: A meta-analytic review. Br J Psychiatry 1995;166(4):424-43. Ineligible outcome format.

219. Pignone M P, Gaynes B N, Rushton J L. Screening for depression in adults: a summary of the evidence for the U.S. Preventive Services Task Force. Ann Intern Med 2002;136:765-76. Review of a non-therapeutic topic.

220. Pirozzo S, Summerbell C, Cameron C et al. Advice on low-fat diets for obesity. Cochrane Database of Systematic Reviews 2002;(2):Not clinical benefit or harm of a specific (class of) drug, device or procedure.

221. Pittler M H, Ernst E. Efficacy of kava extract for treating anxiety: systematic review and metaanalysis. J Clin Psychopharmacol 2000;2084-89. Not clinical benefit or harm of a specific (class of) drug, device or procedure.

222. PORT Meta-analysis, Trialists Group. Postoperative radiotherapy in non-small-cell lung cancer: systematic review and meta-analysis of individual patient data from nine randomised controlled trials. Lancet 1998;352 :257-63. Ineligible MA type.

223. Porthouse J, Cockayne S, King C. Randomised controlled trial of calcium and suplementation with cholecalciferol (vitamin D(subscript 3)) for prevention of fractures in primary care. BMJ 2005;330:1003.Not an SR. 2005;3301003Not an SR.

Abbreviations

AHRQ	Agency for Healthcare Research and Quality
CCJ RCT	Abbreviation use for one of the search strategies tested. CCJ is the Core Clinical Journals, a MEDLINE journal subset previously known Abridged Index Medicus.
CCT	Controlled clinical trial
CQ	Abbreviation use for one of the search strategies tested - Clinical Queries
CR	Abbreviation use for one of the search strategies tested - Citing References
EPC	Evidence-Based Practice Center
HR	Hazard ratio
HTA	Health Technology Assessment
ISI Science Citation	ISI is Institute for Scientific Information
IRQ	Inter-quartile range
MeSH	Medical Subject Headings – the term describes the MEDLINE thesaurus
PMID	PubMed unique identifier
RCT	Randomized controlled trial
RA RCT	Abbreviation use for one of the search strategies tested – Related Article RCT

APPENDIXES

Appendix A. Definitions and Criteria for Signals for Updating

A. Qualitative signals for need to update

We defined signals for two categories of qualitative signals for potential changes in evidence (i.e., for the need to update the original meta-analysis) in terms of their level of importance.

Potentially invalidating change in evidence: one would no longer want clinicians to act upon the results of the original review; an agency or organization that supported the production of the original review would want to retract the review until it could be updated. Examples of such changes include: high quality new evidence that suggests conclusions opposite to those in the original review; high quality new evidence suggests a degree of harm that would completely undermine use of the therapy; or, a head-to-head trial data show that the treatment evaluated in the original review is substantially inferior to another treatment. The specific operational details for each of the three criteria for potentially invalidating changes in evidence are provided below. Importantly, the designation 'potentially invalidating' refers to the recommendations for clinical practice implied by the original meta-analysis, not the methods or conduct of the meta-analysis itself.

Major change in evidence: the conclusions of the original review have not been overturned or superseded, but new evidence clearly has the potential to affect clinical decision-making. Examples of such changes include: new evidence that suggests the therapy does not work in certain patient populations; new evidence that affects how the therapy must be delivered in order to confer the benefit suggested in the original review (e.g., duration of treatment or in conjunction with other co-treatment); evidence about harm that would not completely undermine use of the therapy, but would clearly affect the decision to recommend therapy for at least some patient populations; changes in conclusion that fall short of 'opposite' but to those in the original review; high quality new

Qualitative signals were detected using explicit criteria for comparing the language used to characterize findings in the original meta-analysis with descriptions of findings in new meta-analyses that addressed the same topic, new 'pivotal trials', new clinical practice guidelines, or new editions of major textbooks (e.g., *UpToDate*). Pivotal trials were defined as trials that had a sample size at least three times the previous largest trial or were published in one of the 5 top general medical journals (*New England Journal of Medicine, Lancet, Journal of the American Medical Association, Annals of Internal Medicine,* and the *British Medical Journal*) based on a ranking by journal impact factor. Specific types of qualitative signals are defined below.

Criteria for Signals of Potentially Invalidating Changes in Evidence

A1. Opposing findings: Pivotal trial, meta-analysis including at least one new trial, practice guideline (from major specialty organization or published in peer-reviewed journal), or recent textbook (e.g., *UpToDate*) characterizes the treatment in opposite terms to those in the cohort

review: e.g., definitely effective → ineffective or vice versa (i.e., ineffective → effective). We operationalized 'opposite' as described at the end of this section.

We included guidelines and textbooks as sources of qualitative criteria because our definition of pivotal trial sets a very high bar. For example, we have not included any high impact specialty journals. The only way for a trial not published in one of the top 5 general medical journals to count as a pivotal trial would be for it to have a sample size at least three times that of the previous largest trial. To minimize our overlooking important new evidence while still permitting the efficiency of narrow searches for pivotal trials, we included guidelines and textbooks as sources of qualitative signals for changes in evidence. If new evidence has appeared that is judged of sufficient quality to inform recommendations in practice guidelines or textbooks, then it seems reasonable to call attention to these recommendations as signals for the need to update the original systematic review.

A2. Substantial harm: Pivotal trial, meta-analysis including at least one new trial, practice guideline, recent textbook calls into question the use of the treatment on the basis of harm (i.e., the treatment would no longer be recommended because risks outweigh benefits). A new result for harm that does not undermine use altogether, but has clear potential to affect clinical decision making would count as a 'major change' (criterion A6, 'Important caveat', as defined below).

A3. Superior new treatment: Pivotal trial, systematic review including at least one new trial, practice guideline, or recent textbook characterized another treatment as significantly superior to the one evaluated in the original meta-analysis (based on efficacy or harm)—to the point that it would be preferred in most settings.

Criteria for Signals of Major Changes in Evidence

A4. Important changes in effectiveness short of 'opposing findings': Pivotal trial, new meta-analysis, more recent practice guideline, or recent textbook does not contradict the previous review, but characterizes benefit in substantially different terms (e.g., therapy previously characterized as "promising", "likely beneficial" or similar description and now characterized as definitely beneficial.) This criterion is defined below in greater detail in the explanation of 'Operational definition of changes in conclusions.' Importantly, no attempt was made to distinguish between varying descriptions of "possibly effective." Characterizations such as "may be effective," "promising," "trends towards effectiveness," and other similar phrases or concepts were all categorized as "possibly effective." Thus, this criterion captured substantive differences in the characterization of treatment effects, not merely semantic differences.

A5. Clinically important expansion of treatment: Pivotal trial, new meta-analysis, more recent practice guideline, or recent textbook has expanded of the role of the treatment (e.g., the treatment has now been shown to be of benefit in children or the elderly; or benefit now shown to apply to primary prevention of disease, not just secondary prevention).

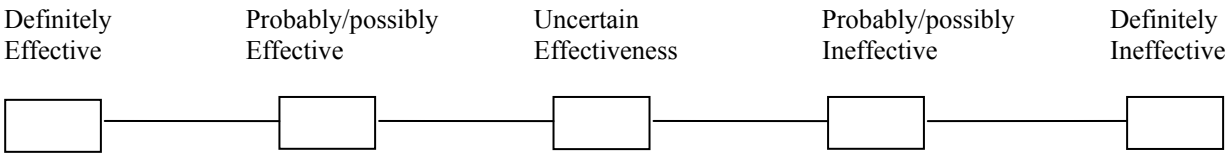
A6. Clinically important caveat: Pivotal trial, new meta-analysis, more recent practice guideline, or recent textbook adds an important caveat, about the patient populations who benefit, way in which treatment has to be delivered in order to derive benefit, sustainability of benefit (e.g., benefits on short term outcomes, but not long-term ones), or increases in harm that are not sufficient to undermine use altogether, but would clearly affect the decision to recommend treatment for at least some patient populations.

A7. Opposing findings from discordant meta-analysis or non-pivotal trial: The treatment has been characterized in sufficiently different terms to the cohort review that disagreement would have met criteria for 'opposing findings' (criterion A1) except the source was not a pivotal trial, new-meta-analysis, or more recent practice guideline, or recent textbook—rather, it was a discordant meta-analysis or trial indexed in *ACP Journal Club*. ('Discordant meta-analysis' was defined as one that reached different conclusions than the original meta-analysis, despite effectively covering the same search period.)

We included this criterion because our definition of pivotal trial sets a very high bar, including only the top 5 general medical journals and trials with sample sizes at least three times the size of the previous largest trial. This criterion allows other sources of evidence to count as qualitative singles, without allowing any new trial with different results than in the previous systematic review to count as a signal for updating.

Operational definition of changes in conclusions

Labels such as 'effective' and 'ineffective' do not capture the distinction between trends towards effectiveness or uncertainty in the face of conflicting results or major limitations of the existing evidence. On the other hand, attempting to capture such nuances runs the risk of regarding semantic or stylistic differences between different authors. To balance these concerns, we consider conclusions about effectiveness in terms of a 5-point scale as shown below.



For systematic reviews that focused on adverse effects of treatment, we replaced effective/ineffective with 'harmful/not harmful.

In the interest of having qualitative signals of changes in evidence with high specificity, we did not attempt to make distinctions between statements of 'probable' or 'possible' benefit (or lack of benefit). We assigned descriptions such as 'promising,' 'possibly', 'probably', 'maybe', 'likely' into the same category. While still subjective, our labels are thus quite conservative—we are distinguishing firm or confident results, from trends, and equipoise or complete uncertainty (the middle position).

We defined 'opposite' conclusions (criterion A1 for potentially invalidating changes in evidence) as a movement of at least two positions on the above scale and 'important changes in effectiveness short of opposing findings' (criterion A4 for major changes in evidence) as a movement of one position on this scale. A movement of two positions generally includes movements from benefit to lack of benefit (or vice versa), but also includes movements from uncertain to definite conclusions about effectiveness. In this context, it is important to emphasize that we were careful not to equate summary conclusions (e.g., in article abstracts) of the type "the evidence does not permit definite conclusions" with 'complete uncertainty.' In many such

cases, the results reported in the trial or meta-analysis indicated a trend, but the authors regarded the trend as inconclusive, on statistical or methodological grounds. Such cases were judged as 'possible' benefit (or lack of benefit, depending on the results). We reserved 'completely uncertain' for cases in which the authors clearly regarded the evidence as not indicating towards either benefit or lack of benefit. Thus, we regarded that a change from a definite or confident conclusion to compete uncertainty (or vice versa) would represent a potentially invalidating change in evidence. For example this would include a change from there being no basis on which to recommend a treatment to its being definitely recommended (or vice versa).

B. Quantitative signals of changes in evidence

We performed updated meta-analyses that combined the results from new trials with the meta-analytic result reported in the original review. To count as a quantitative signal, an outcome explicitly identified as a primary outcome in the original meta-analysis or any mortality outcome had to meet one of the criteria below. To count as a primary outcome, we required use of the word 'primary' or 'main.' Even in cases where those words were used, we discounted such outcomes if authors stated that they had more than 3 such outcomes (on the grounds that more than 3 undermines the concept of 'primary').

B1. Change in statistical significance: at least one of the 95% confidence limits lies on a different side of the line of no effect (i.e. odds ratio or relative risk=1, risk difference=0). This criterion captures whether a result that was statistically significant in the original systematic review is now not statistically significant or vice versa—a previously non-significant result has become statistically significant.

To avoid counting trivial or 'borderline' changes in statistical significance as quantitative signals for updating, we required that at least one of the two results (i.e., the original and updated meta-analyses) have a p-value outside the range of 0.04 to 0.06. In other words, we excluded cases in which the original systematic review reported a borderline result and the updated result is also borderline but happens to lie in the other side of the line of no effect. For instance, a change from p =0.041 to p =0.059 would not count as a quantitative signal to update, nor would the converse change (from p =0.059 to p =0.041).

B2: Change in effect size of at least 50%: the new result indicates a relative change in effect size of at least 50%. For example, if RRRnew / RRRold <=0.5 or RRRnew / RRRold >=1.5, where RRR is the relative risk reduction. Thus, if the original review has found RR=0.70 for mortality, this implies RRR of 0.3. If the updated meta-analytic result for mortality were 0.90, then the updated RRR would be 0.10, which is less than 50% of the previous RRR. In other words the reduction in the risk of death has moved from 30% to 10%. The same criterion applied for odds ratios (e.g., if previous OR=0.70 and updated result were OR=0.90, then the new reduction in odds of death (0.10) would be less 50% of the magnitude of the previous reduction in odds (0.30). For risk differences and weighted mean differences, we applied the criterion directly to the previous and updated results (e.g., RDnew / RDold <=0.5 or RDnew / RDold >=1.5).

Appendix B: Examples of Systematic Reviews with Qualitative Signals for Potentially Invalidating Changes in Evidence

I. All 8 Reviews with Signals for "Potentially Invalidating Changes in Evidence" (criteria for signals A1-A3)

---

1. **The Albumin Reviewers (Alderson P, Bunn F, Lefebvre C, Li Wan Po A, Li L, Roberts I, Schierhout G). Human albumin solution for resuscitation and volume expansion in critically ill patients (Cochrane Review). In: *The Cochrane Library*, Issue 2, 2001. Oxford: Update Software.**

Question(s) addressed	Does human albumin or plasma protein fraction reduce mortality in patients who are critically ill (hypovolemia, burns, or hypoalbuminemia)?
Findings of Original Review	The original review found that "For each patient category the risk of death in the albumin treated group was higher than in the comparison group… an increase in the risk of death of 6% (3% to 9%). These data suggest that for every 17 critically ill patients treated with albumin there is one additional death."
New Findings	A pivotal trial found no difference in the risk of death between patients who received albumin and those who did not (relative risk of death, 0.99; 95 percent confidence interval, 0.91 to 1.09; p =0.87). It concluded: "In patients in the ICU, use of either 4 percent albumin or normal saline for fluid resuscitation results in similar outcomes at 28 days."
Qualitative signal	**Opposing findings**: The original systematic review reported an increase in mortality; a pivotal trial showed no difference in risk of death
Quantitative signal	**Change in statistical significance**: Relative risk of death became non-significant RR = 1.68 (1.26, 2.23) → 1.04 (0.95, 1.13) **Change in effect magnitude of 50% or more**: Relative risk increase for death of 0.68 → increase of only 0.04
Other signals	Increase in number of patients of at least 50%: N=1419 → N=8352 Trial with sample size at least 3 times the size of previous largest trial: Previous largest trial had 219 patients; new trial had 6933 patients Change in width of 95% confidence interval of at least 50%: as shown above
Source of	**Pivotal Trial**

new evidence	Finfer S et al. A comparison of albumin and saline for fluid resuscitation in the intensive care unit. N Engl J Med. 2004;350:2247-56.
Time to signal	Qualitative signal: 3.0 years Quantitative signal: same

---

2. **Alejandria MM, Lansang MA, Dans LF, Mantaring JBV. Intravenous immunoglobulin for treating sepsis and septic shock (Cochrane Review). In: *The Cochrane Library*, Issue 1, 2001. Oxford: Update Software.**

Question(s) addressed	Does intravenous immunoglobulin (IVIG) reduce mortality, bacteriological failure rates, and duration of stay in hospital in patients with bacterial sepsis septic shock?
Findings of Original Review	Comparing polyclonal IVIG versus control, the original review reported a relative risk of death 0.60 (95% CI: 0.47 to 0.76) among a total of 413 patients. The authors concluded that polyclonal intravenous immunoglobulin "significantly reduces mortality and can be used as an adjuvant treatment for sepsis and septic shock."
New Findings	A subsequent meta-analysis (Pildal 2004) included 763 patients and found that "[h]igh-quality trials …showed a relative risk of 1.02 (95% CI, 0.84-1.24), whereas other trials (involving a total of 948 patients, 292 of whom died) showed a relative risk of 0.61 (95% CI, 0.50-0.73). Because high-quality trials failed to demonstrate a reduction in mortality, polyclonal immunoglobulin should not be used for treatment of sepsis except in randomized clinical trials." The textbook Up-To-Date quotes this subsequent meta-analysis and states intravenous immunoglobulin "is rarely used to treat patients with septic shock in the United States, and this approach is not recommended pending the demonstration of benefit in large, well designed trials."
Qualitative signal	**Opposing findings**: The original systematic review reported a definite reduction in mortality; a subsequent meta-analysis showed no benefit
Quantitative signal	**Change in statistical significance:** among higher quality trials only **Change in effect magnitude of 50% or more**: among higher quality trials only
Other signals:	Increase in number of patients of at least 50%: N=1992 → N=3082 (this increase was for meta-analysis of monoclonal anti-endotoxins; for polyclonal IVIG, increase was not 47%)
Source(s) of new evidence	Pildal J, Gotzsche PC. Polyclonal immunoglobulin for treatment of bacterial sepsis: a systematic review. Clin Infect Dis. 2004;39(1):38-46.

Time to signal	Qualitative signal: 3.0 years Quantitative signal: not applicable

---

3. **Bucher, H. C., Guyatt, G. H., and Cook, R. J., Effect of calcium supplementation on pregnancy-induced hypertension and preeclampsia: A meta-analysis of randomized controlled trials. *JAMA.* 1996. 275: 1113-1117.**

Question(s) addressed?	What effects does calcium supplementation during pregnancy have on blood pressure, preeclampsia, and adverse maternal and fetal outcomes
Findings of Original Review	The original review showed a substantial, statistically significant reduction in the occurrence of preeclampsia among women who received calcium supplementation compared with placebo was (OR of 0.38; 95% CI, 0.22 to 0.65), as well as significant improvements in blood pressure. It concluded: "Calcium supplementation during pregnancy leads to an important reduction in systolic and diastolic blood pressure and preeclampsia."
New Findings	A pivotal trial published the following year reported: "Calcium supplementation did not significantly reduce the incidence or severity of preeclampsia or delay its onset… There were no significant differences between the two groups in the prevalence of pregnancy-associated hypertension without preeclampsia (15.3 percent vs. 17.3 percent) or of all hypertensive disorders (22.2 percent vs. 24.6 percent). The mean systolic and diastolic blood pressures during pregnancy were similar in both groups." It concluded that "Calcium supplementation during pregnancy did not prevent preeclampsia, pregnancy-associated hypertension, or adverse perinatal outcomes in healthy nulliparous women."
Qualitative signal	**Opposing findings**: The original systematic review reported a definite reduction in pre-eclampsia and development of hypertension; a pivotal trial showed no impact on either outcome.
Quantitative signal	**Change in effect magnitude of 50% or more**: reduction in odds of pre-eclampsia of 0.62 → reduction of only 0.21; and reduction in odds of developing hypertension of 0.70 → reduction of only 0.25
Other signals:	Increase in number of patients of at least 50%: N=2280→ N=7059 Trial with sample size at least 3 times the size of previous largest trial: previous largest trial included 1167 patients; new trial included 4779 patients
Source(s) of new evidence	**Pivotal trial**: Levine RJ, Hauth JC, Curet LB, et al. Trial of calcium to prevent preeclampsia. N Engl J Med. 1997;337:69-76.

Additional trials:

1. Cong K et al. Calcium supplementation during pregnancy for reducing pregnancy induced hypertension. Chin Med J (Engl) 1995;108:57-9.
2. Purwar M et al. Calcium supplementation and prevention of pregnancy induced hypertension. J Obstet Gynaecol Res. 1996;22:425-30.

Time to signal Qualitative signal: 1.2 years
Quantitative signal: same

---

4. **Coward LJ et al. Safety and efficacy of endovascular treatment of carotid artery stenosis compared with carotid endarterectomy: a Cochrane systematic review of the randomized evidence. *Stroke*. 2005;36(4):905-11.**

Question(s) addressed In patients with carotid stenosis, what are the risks and benefits of endovascular treatment compared with carotid endarterectomy?

Findings of Original Review The original review found no significant difference in the odds of treatment related death or any stroke (odds ratio [OR], endovascular surgery, 1.33; 95% confidence interval [CI], 0.86 to 2.04), death or disabling stroke (OR, 1.22; CI, 0.61 to 2.41), or death, any stroke, or myocardial infarction (OR, 1.04; CI, 0.69 to 1.57). At 1 year after randomization, there was no significant difference between the 2 treatments in the rate of any stroke or death (OR, 1.01; CI, 0.71 to 1.44).

It concluded: "No significant difference in the major risks of treatment was found but the wide confidence intervals indicate that it is not possible to exclude a difference in favor of one treatment. Minor complication rates favor endovascular treatment."

New Findings One pivotal trial (Mas 2006) was stopped early because of significantly inferior outcomes for endovascular treatment. "The 30-day incidence of any stroke or death was 3.9% after endarterectomy (95% CI: 2.0 to 7.2) and 9.6% after stenting (95% CI: 6.4 to 14.0); the relative risk of any stroke or death after stenting as compared with endarterectomy was 2.5 (95% CI: 1.2 to 5.1)." Rates of death and stroke at 6 months were also lower with endarterectomy than with stenting.

Another pivotal trial (Ringleb 2006) found that "The rate of death or ipsilateral ischemic stroke from randomization to 30 days after the procedure was 6.84% with carotid-artery stenting and 6.34% with carotid endarterectomy (absolute difference 0.51%, 90% CI –1.89% to 2.91%). Based on a pre-defined non-inferiority margin of 2.5%, the authors concluded that endovascular treatment "failed to prove non-inferiority of carotid-artery stenting compared with carotid endarterectomy… The results of this trial do

not justify the widespread use in the short-term of carotid-artery stenting for treatment of carotid-artery stenoses."

Qualitative signal

Opposing findings: The original review reported no major differences between the two treatments. The review emphasized the uncertainty of the comparison, but did not specifically indicate any possibility that endovascular treat was inferior to endarterectomy. Two pivotal trials indicate inferiority of endovascular treatment (in one case, of sufficient magnitude to result in termination of the trial).

Editorials for both pivotal trials discuss possible explanation for these findings that leave open the possibility of non-inferiority. But the point remains that the publication of these two high profile trials with results substantially different from those of previous trials constitutes an important signal for the need for updating the original systematic review.

Quantitative signal

Change in statistical significance: Relative risk of stroke or death within 30 days became statistically significant, with both limits of 95% confidence interval now lying on side of increased risk with endovascular treatment

RR = 1.33 (0.86, 2.04)→ 1.35 (1.02, 1.80)

Other signals:

Increase in number of trials of at least 50%: 6 trials → 9 trials
Increase in number of patients of at least 50%: N=1269 → 3376

Source(s) of new evidence

Pivotal trials:

1. Mas JL, Chatellier G, Beyssen B, et al. Endarterectomy versus stenting in patients with symptomatic severe carotid stenosis. N Engl J Med. 2006;355:1660-71.
2. Ringleb PA, Allenberg J, Bruckmann H, et al. 30 day results from the SPACE trial of stent-protected angioplasty versus carotid endarterectomy in symptomatic patients: a randomized non-inferiority trial. Lancet. 2006;368:1239-47

Additional trial:
Carotid Revascularization Using Endarterectomy or Stenting Systems (CaRESS) phase I clinical trial: 1-year results. J Vasc Surg. 2005;42:213-

Time to signal

Qualitative signal: 1.5 years
Quantitative signal: same

---

5. **Hemminki E, McPherson K. Impact of postmenopausal hormone therapy on cardiovascular events and cancer: pooled data from clinical trials. *BMJ*. 1997;315:149-53.**

Question(s) addressed	Is hormone replacement therapy (HRT) associated with cardiovascular events or cancer in postmenopausal women?
Findings of Original Review	The original review concluded that was no clear evidence of an association between cardiovascular outcomes and HRT, but noted that "Data on cardiovascular events and cancer were usually given incidentally, either as a reason for dropping out of a study or in a list of adverse effects." We therefore characterized the original systematic review as having concluded that effectiveness was 'uncertain'.
New Findings	A pivotal trial (Hulley 1998) found no difference between HRT and placebo in terms of the primary or secondary cardiovascular endpoints. (RR=0.99; 95% CI: 0.80 to 1.22). The trial also showed an increase in thromboembolic events. It concluded: "Based on the finding of no overall cardiovascular benefit and a pattern of early increase in risk of CHD events, we do not recommend starting this treatment for the purpose of secondary prevention of CHD." A second, larger pivotal trial (Rossouw 2002) was stopped early "because the test statistic for invasive breast cancer exceeded the stopping boundary for this adverse effect and the global index statistic supported risks exceeding benefits." Based on a mean follow-up of 5.2 years, "[a]bsolute excess risks per 10 000 person-years attributable to estrogen plus progestin [HRT] were 7 more CHD events, 8 more strokes, 8 more PEs, and 8 more invasive breast cancers, while absolute risk reductions per 10 000 person-years were 6 fewer colorectal cancers and 5 fewer hip fractures. The absolute excess risk of events included in the global index was 19 per 10 000 person-years. "Overall health risks exceeded benefits from use of combined estrogen plus progestin for an average 5.2-year follow-up among healthy postmenopausal US women. All-cause mortality was not affected during the trial. The risk-benefit profile found in this trial is not consistent with the requirements for a viable intervention for primary prevention of chronic diseases, and the results indicate that this regimen should not be initiated or continued for primary prevention of CHD."
Qualitative signal	**Opposing findings:** The original systematic review found no clear relationship between HRT and cardiovascular outcomes. Two pivotal trials clearly demonstrated a lack of benefit and evidence of some harm.
Quantitative signal	**Change in statistical significance**: odds of *increased* cardiovascular and thromboembolic events became statistically significant.

	Odds ratio of 1.64 (0.65, 4.18) → 1.70 (1.18, 2.43)
Other signals:	Increase in number of patients of at least 50%: N=4124 → 25140
	Trial with sample size at least 3 times the size of previous largest trial: previous largest trial had N=1265; new trial had N=16608
	Change in width of 95% confidence interval of at least 50%: as shown above
Source(s) of new evidence	**Pivotal trials**:

1. Hulley S, Grady D, Bush T, et al. Randomized trial of estrogen plus progestin for secondary prevention of coronary heart disease in postmenopausal women. Heart and Estrogen/progestin Replacement Study (HERS) Research Group. JAMA. 1998;280(7):605-13.
2. Herrington DM, Reboussin DM, Brosnihan KB, et al. Effects of estrogen replacement on the progression of coronary-artery atherosclerosis. N Engl J Med. 2000;343(8):522-9.
3. Waters DD, Alderman EL, Hsia J, et al. Effects of hormone replacement therapy and antioxidant vitamin supplements on coronary atherosclerosis in postmenopausal women: a randomized controlled trial. JAMA. 2002;288(19):2432-40.
4. Rossouw JE, Anderson GL, Prentice RL, et al. Risks and benefits of estrogen plus progestin in healthy postmenopausal women: principal results From the Women's Health Initiative randomized controlled trial. JAMA.2002;288(3):321-33.

Time to signal	Qualitative signal: 1.1 years Quantitative signal: same

---

6. **Kjaergard LL, Krogsgaard K, Gluud C. Interferon alfa with or without ribavirin for chronic hepatitis C: systematic review of randomized trials. *BMJ*. 2001;323:1151-5.**

Question(s) addressed	How efficacious and safe is interferon alfa with or without ribavirin in the treatment of chronic hepatitis C?
Findings of Original Review	The original review found that, compared with interferon alone, "combination therapy reduced the risk of not having a sustained virological for 6 months by 26% in naïve patients (relative risk 0.74, 95% confidence interval 0.70 to 0.78), 33% in relapsers (0.67, 0.57 to 0.78), and 11% in non-responders (0.89, 0.83 to 0.96). Morbidity and mortality showed a non-significant trend in favour of combination therapy (Peto odds ratio 0.45, 0.19 to 1.06). Combination therapy significantly reduced the risk of not having improvement in results of histology by 17% in naive patients (0.83, 0.74 to 0.93) and by 27% in relapsers and non-responders (0.73, 0.66 to 0.82). The authors concluded that "treatment with interferon alfa plus

ribavirin has a significant beneficial effect on the virological and histological responses of patients with chronic hepatitis C…"

New Findings Two pivotal trials compared the combination evaluated in the original systematic review with an alternative treatment, peginterferon alfa combined with ribavirin.

The first trial included three treatment arms, standard interferon alfa-2b plus ribavirin (as evaluated in the original review), pegylated interferon alfa-2b (1.5 µg/kg per week for four weeks followed by 0.5 µg/kg per week) plus ribavirin, and pegylated interferon alfa-2b (1.5 µg/kg per week) plus ribavirin. The primary endpoint of sustained virologic response "was significantly higher ($p = 0.01$ for both comparisons) in the higher-dose peginterferon group (274/511 [54%]) than in the lower-dose peginterferon (244/514 [47%]) or interferon (235/505 [47%]) groups."

They concluded "In patients with chronic hepatitis C, the most effective therapy is the combination of peginterferon alfa-2b 1.5 microg/kg per week plus ribavirin," though they noted that "The benefit is mostly achieved in patients with HCV genotype 1 infections."

The second pivotal trial (Fried 2002) found that "a significantly higher proportion of patients who received peginterferon alfa-2a plus ribavirin had a sustained virologic response (defined as the absence of detectable HCV RNA 24 weeks after cessation of therapy) than of patients who received interferon alfa-2b plus ribavirin (56 percent vs. 44 percent, $p < 0.001$) or peginterferon alfa-2a alone (56 percent vs. 29 percent, $p < 0.001$)." They concluded: "In patients with chronic hepatitis C, once-weekly peginterferon alfa-2a plus ribavirin was tolerated as well as interferon alfa-2b plus ribavirin and produced significant improvements in the rate of sustained virologic response, as compared with interferon alfa-2b plus ribavirin or peginterferon alfa-2a alone."

The textbook Up-To-Date cites these two trials (and a subsequent trial that evaluated the optimal doe of ribavarin) in making the statement that "combination therapy with pegylated interferon plus ribavirin is generally associated with a higher sustained virologic response rate compared to combination therapy with standard interferon plus ribavirin or pegylated interferon monotherapy. As a result, this is usually the preferred approach in patients with hepatitis C who have not previously received treatment." The chapter in Up-To-Date noted the influence of genotype on response, which was seen in both trials. Because the benefit in the first trial was largely confined to patients with a particular genotype, we did not take that trial by itself as the basis for the signal of a superior alternate treatment. We regarded the signal as triggered by the second trial (Fried 2002).

Qualitative **Superior new treatment**: Head to head comparisons in two pivotal trials

signal	showed that an alternative treatment is superior to the therapy evaluated in the original systematic review.
Quantitative signal	Not applicable – comparisons in new trials differ from those in the original systematic review

Other signals: None

Source(s) of new evidence **Pivotal trials**:

1. Manns MP, McHutchison JG, Gordon SC, et al. Peginterferon alfa-2b plus ribavirin compared with interferon alfa-2b plus ribavirin for initial treatment of chronic hepatitis C: a randomized trial. Lancet. 2001;358:958-65.
2. Fried MW, Shiffman ML, Reddy KR, et al. Peginterferon alfa-2a plus ribavirin for chronic hepatitis C virus infection. N Engl J Med. 2002;347(13):975-82.

Time to signal Qualitative signal: 0.9 years
Quantitative signal: Not applicable

---

7. **Lefering R, Neugebauer EA. Steroid controversy in sepsis and septic shock: a meta-analysis. *Crit Care Med*. 1995;23(7):1294-303.**

Question(s) addressed	Does the use of corticosteroids in patients with sepsis or septic shock lower the risk of death?
Findings of Original Review	The original review found that "Corticosteroids did not change 28 day mortality (15 trials, n = 2022; relative risk 0.92, 95% confidence interval 0.75 to 1.14) or hospital mortality (13 trials, n = 1418; 0.89, 0.71 to 1.11)." The authors concluded that "No overall beneficial effect of corticosteroids in patients with septic shock was observed…"
New Findings	A randomized, double-blind, multi-center trial evaluated the impact of a 7-day course of low-dose hydrocortisone versus placebo in patients who showed signs of relative adrenal insufficiency. It found a significantly lower risk of death in the corticosteroid group (hazard ratio, 0.67; 95% confidence interval, 0.47-0.95; p =0.02). It concluded that "a 7-day treatment with low doses of hydrocortisone and fludrocortisone significantly reduced the risk of death in patients with septic shock and relative adrenal insufficiency without increasing adverse events." A subsequent meta-analysis showed that, among five trials (n = 465) involving long courses (> or = 5 days) with low dose (< or = 300 mg hydrocortisone or equivalent), the relative risk for mortality at 28 days was 0.80 (95% CI: 0.67 to 0.95).

Qualitative signal	**Opposing findings**: The original systematic review found no mortality benefit regardless of dose. A pivotal trial and subsequent meta-analysis showed define reductions in mortality with low dose regimens given for at least 5 days.
Quantitative signal	**Change in effect magnitude of 50% or more**: the absolute risk reduction for mortality increased from 0.2% to 4% (the criterion was first met at after Slusher 1996, when updated risk reduction increased to 1%)
Other signals	Increase in number of patients of at least 50%: N=530 → N=1067 Increase in number of trials of at least 50%: 10 trials → 16 trials Change in width of 95% confidence interval of at least 50%: The original 95% CI for mortality with low-dose steroids extended from a 20% absolute reduction to a 16% increase in mortality. The 95% CI for the updated result extended from a 13% reduction to a 1% increase.
Source(s) of new evidence	**Pivotal trial**: Annane D, Sebille V, Charpentier C, et al. Effect of treatment with low doses of hydrocortisone and fludrocortisone on mortality in patients with septic shock. JAMA. 2002;288:862-71. **Additional trials and meta-analysis**: 1. Bollaert PE, Charpentier C, Levy B, Debouverie M, Audibert G, Larcan A. Reversal of late septic shock with supraphysiologic doses of hydrocortisone. Crit Care Med. 1998;26(4):645-50. 2. Briegel J, Forst H, Haller M, et al. Stress doses of hydrocortisone reverse hyperdynamic septic shock: a prospective, randomized, double-blind, single-center study. Crit Care Med. 1999;27(4):723-32. 3. Slusher T, Gbadero D, Howard C, et al. Randomized, placebo-controlled, double blinded trial of dexamethasone in African children with sepsis. Pediatr Infect Dis J. 1996;15(7):579-83. 4. Annane D, Bellissant E, Bollaert PE, Briegel J, Keh D, Kupfer Y. Corticosteroids for severe sepsis and septic shock: a systematic review and meta-analysis. BMJ. 2004;329(7464):480.
Time to signal	Qualitative signal: 7.1 years Quantitative signal: 1 year

---

8. **Lord JM, Flight IH, Norman RJ. Metformin in polycystic ovary syndrome: systematic review and meta-analysis. *BMJ*. 2003;327:951-953.**

Question(s) addressed	Does metformin improve pregnancy and ovulation rates in women with polycystic ovary syndrome?

Findings of Original Review	The original review found that "metformin is effective in achieving ovulation in women with polycystic ovary syndrome, with odds ratios of 3.88 (95% confidence interval 2.25 to 6.69) for metformin compared with placebo and 4.41 (2.37 to 8.22) for metformin and clomifene compared with clomifene alone. An analysis of pregnancy rates shows a significant treatment effect for metformin and clomifene (odds ratio 4.40, 1.96 to 9.85)." Referring to the use of metformin, the authors concluded "its choice as a first line agent seems justified."
New Findings	A pivotal trial compared clomifene citrate plus metformin with clomifene plus placebo and found a lower ovulation rate in the metformin group "(64% compared with 72% in the placebo group, a non-significant difference (risk difference − 8%, 95% confidence interval − 20% to 4%). There were no significant differences in either rate of ongoing pregnancy (40% v 46%; − 6%, − 20% to 7%) or rate of spontaneous abortion (12% v 11%; 1%, − 7% to 10%). A significantly larger proportion of women in the metformin group discontinued treatment because of side effects (16% v 5%; 11%, 5% to 16%)." The authors concluded that "metformin is not an effective addition to clomifene citrate as the primary method of inducing ovulation in women with polycystic ovary syndrome." The accompanying editorial also concluded that "metformin should not be used routinely as part of first line treatment for inducing ovulation."
Qualitative signal	**Opposing findings**: The original systematic review concluded that metformin is definitely effective, recommending it as a first line agent. A pivotal trial showed no benefit and concluded that metformin should not be considered a first line treatment.
Quantitative signal	**Change in statistical significance**: increase in ovulation rate in patients treated with metformin and clomifene vs. clomifene alone lost statistical significance Odds ratio of 4.41 (2.37, 8.22) → 1.42 (0.98, 2.05) **Change in effect magnitude of 50% or more**: Relative increase in ovulation rate in patients treated with metformin and clomifene vs. clomifene alone decreased by over 50% (OR of 4.41→ 1.42), as did the relative increase in clinical pregnancy rate among patients who received metformin and clomifene vs. clomifene alone (OR of 4.40→ 2.07)
Other signals:	Increase in number of patients of at least 50%: for the outcome of clinical pregnancy rate, the number of patients increased from 173 to 537 Increase in number of trials of at least 50%: for the outcome of clinical pregnancy rate, the number of trials increased from 3 to 8

	Change in width of 95% confidence interval of at least 50%: as shown above
Source(s) of new evidence	**Pivotal trial**: Moll E et al. Effect of clomifene citrate plus metformin and clomifene citrate plus placebo on induction of ovulation in women with newly diagnosed polycystic ovary syndrome: randomized double blind clinical trial. BMJ 2006;332:1485. Four additional trials contained in meta-analysis: Kashyap S, Wells GA, Rosenwaks Z. Insulin-sensitizing agents as primary therapy for patients with polycystic ovarian syndrome. Hum Reprod. 2004;19:2474-83.
Time to signal	Qualitative signal: 2.6 years Quantitative signal: same

---

II. Examples of Reviews with Signals for "Major Changes in Evidence" (criteria A4-A7)

Examples of criterion A4: Pivotal trial, new meta-analysis, more recent practice guideline, or recent textbook does not contradict the previous review, but characterizes benefit in substantially different terms (e.g., therapy previously characterized as "promising", "likely beneficial" or similar description and now characterized as definitely beneficial.)

Original Review	Collaborative meta-analysis of randomised trials of antiplatelet therapy for prevention of death, myocardial infarction, and stroke in high risk patients. *BMJ*. 2002;324(7329):71-86.
Question(s) addressed	Covered a variety of questions related to the effects of antiplatelet therapy among patients at high risk of occlusive vascular events, including: Is aspirin plus dipyridamole was more effective than aspirin alone for the secondary prevention of vascular events after ischemic stroke of presumed arterial origin?
Findings of Original Review	The original review stated that "the addition of dipyridamole to aspirin was associated with only a non-significant further 6% (6%) reduction in serious vascular events…The apparent reduction in non-fatal stroke was derived mainly from one large study… but this result was not supported by the findings for non-fatal stroke in the other studies." It concluded: "Addition of dipyridamole to aspirin produced no significant further reduction in vascular events compared with aspirin alone."

New Findings A pivotal trial found that patients who received aspirin and dipyridamole had a significantly lower risk of the primary outcome (a composite of death from all vascular causes, non-fatal stroke, non-fatal myocardial infarction, or major bleeding complication, whichever happened first), with a hazard ratio 0.80, 95% CI 0.66-0.98; absolute risk reduction 1% per year, 95% CI 0.1-1.8). Combining these data with previous trials resulted in an overall risk ratio for the composite of vascular death, stroke, or myocardial infarction of 0.82 (95% CI 0.74-0.91). The authors concluded: "The ESPRIT results, combined with the results of previous trials, provide sufficient evidence to prefer the combination regimen of aspirin plus dipyridamole over aspirin alone as antithrombotic therapy after cerebral ischaemia of arterial origin."

Qualitative signal **Major change**: possibly superior → definitely superior

Quantitative signal **Change in statistical significance**: The lower risk of serious vascular events (vascular death or death from unknown cause, MI or stroke) became statistically significant.

Odds ratio of 0.94 (0.84, 1.06) → 0.90 (0.81, 0.99)

As noted above, the random effects meta-analytic result for relative risk is 0.82 (0.74-0.91), which more clearly shows the change in statistical significance. Odds ratios were used in our analysis because the original review used odds ratios.

Other signals Because the original review covered a number of distinct questions related to antiplatelet therapy for the prevention of vascular events, other qualitative and quantitative and signals may have been met. For example, a pivotal trial found that adding aspirin to clopidogrel increased bleeding without reducing recurrent ischemic vascular events in high-risk patients.[2] Another pivotal trial found that clopidogrel plus aspirin did not differ from aspirin alone for reducing MI, stroke, and cardiovascular death in patients with clinically evident cardiovascular disease or multiple risk factors.[3]

Both of these qualitative signals occurred prior to the signal involving the comparison of aspirin plus dipyridamole with aspirin alone, but the latter more clearly fit one of our qualitative criteria and involved a quantitative signal as well.

Source(s) of **Pivotal trial:**

evidence Halkes PH, van Gijn J, Kappelle LJ, Koudstaal PJ, Algra A. Aspirin plus dipyridamole versus aspirin alone after cerebral ischaemia of arterial origin (ESPRIT): randomised controlled trial. Lancet. 2006;367(9523):1665-73.

Additional pivotal trials addressing other questions in the original review:

1. Diener HC, Bogousslavsky J, Brass LM, et al. Aspirin and clopidogrel compared with clopidogrel alone after recent ischaemic stroke or transient ischaemic attack in high-risk patients (MATCH): randomised, double-blind, placebo-controlled trial. Lancet. 2004;364(9431):331-7.
2. Bhatt DL, Fox KA, Hacke W, et al. Clopidogrel and aspirin versus aspirin alone for the prevention of atherothrombotic events. N Engl J Med. 2006;354(16):1706-17.

Time to signal Qualitative signal: 4.4 years
Quantitative signal: same

Original Review Avezum A, Tsuyuki RT, Pogue J, Yusuf S. Beta-blocker therapy for congestive heart failure: a systemic overview and critical appraisal of the published trials. Can J Cardiol. 1998;14(8):1045-53.

Question(s) addressed Do beta-blockers reduce mortality and morbidity in the treatment of heart failure?

Findings of Original Review The original review reported a lower odds of death with beta-blockers that had borderline statistical significance (OR = 72; 99% CI 0.51 to 1.00). The authors were concerned about the sparseness of the data on mortality compared with evaluations of beta-blockers of patients with myocardial infarction. They concluded: "Although the effects on mortality were nominally statistically significant, the use of formal methods of interim monitoring adapted for meta-analyses suggests that substantially more patients still need to be studied in large scales trials to provide reliable and conclusive evidence."

New Findings A pivotal trial (MERIT-HF 1999) was stopped early because of the magnitude of reduction in the beta-blocker group, with a relative risk 0.66 (95% CI 0.53-0.81; p =0.00009 or adjusted for interim analyses p =0.0062). The authors concluded: "Metoprolol CR/XL once daily in addition to optimum standard therapy improved survival." A second pivotal trial (CIBIS-II 1999) published the same year was also stopped early because of the survival benefit evident in the beta-blocker group. A third pivotal trial (Packer 2001) demonstrated a significant reduction in mortality for patients with more severe heart failure.

Qualitative Major change signal: possible mortality benefit → definite benefit

Quantitative signal **Change in statistical significance: borderline** reduction in mortality became statistically significant

0.72 (0.51, 1.00) → 0.67 (0.49, 0.91)

This change reflects the first shift to statistical significance (after Herlitz 1997); after additional trials, the updated result was 0.78 (0.70, 0.88)

Other signals: Increase in number of trials of at least 50%: 10 trials → 15 trials

Increase in number of patients of at least 50%: N= 2841 → N=14738

Trial with sample size at least 3 times the size of previous largest trial: previous largest trial included 1094 patients; a new trial included 3991 patients

Source(s) of evidence **Pivotal trial:**
Effect of metoprolol CR/XL in chronic heart failure: Metoprolol CR/XL Randomised Intervention Trial in Congestive Heart Failure (MERIT HF). *Lancet*. 1999;353(9169):2001-7.

Additional trials (including two pivotal trials):

1. Herlitz J, Waagstein F, Lindqvist J, Swedberg K, Hjalmarson A. Effect of metoprolol on the prognosis for patients with suspected acute myocardial infarction and indirect signs of congestive heart failure (a subgroup analysis of the Goteborg Metoprolol Trial). *Am J Cardiol*. 1997;80(9B):40J-44J.
2. The Cardiac Insufficiency Bisoprolol Study II (CIBIS-II): a randomised trial. *Lancet*. 1999;353(9146):9-13.
3. Packer M, Coats AJ, Fowler MB, et al. Effect of carvedilol on survival in severe chronic heart failure. *N Engl J Med*. 2001;344(22):1651-8.

Time to signal Qualitative signal: 1 year
Quantitative signal: -0.6 years

Original Review Birck R, Krzossok S, Markowetz F, Schnulle P, van der Woude FJ, Braun C. Acetylcysteine for prevention of contrast nephropathy: meta-analysis. *Lancet* 2003;362:598-603.

Question(s) addressed Does prophylactic acetylcysteine reduces contrast nephropathy in patients with chronic renal insufficiency?

Findings of Original Review The original review included 7 trials and found that "compared with periprocedural hydration alone, administration of acetylcysteine and hydration significantly reduced the relative risk of contrast nephropathy by 56% (0.435

[95% CI 0.215-0.879], p =0.02) in patients with chronic renal insufficiency. Meta- regression revealed no significant relation between the relative risk of contrast nephropathy and the volume of radiocontrast media administered or the degree of chronic renal insufficiency before the procedure." The authors acknowledged that it remained unclear to what extent acetylcysteine improved harder clinical endpoints, but the impact on measures of renal function was regarded as robust. They concluded "acetylcysteine with hydration significantly reduces the risk of contrast nephropathy in patients with chronic renal insufficiency."

New Findings A subsequent meta-analysis (published 1.4 years after the first) included 20 trials and found that the impact on contrast nephropathy was smaller in magnitude and of borderline statistical significance. The authors also emphasized that the trials showed significant heterogeneity that remained unexplained despite exploration of various possible clinical and methodological differences across the studies.

They concluded: "Acetylcysteine may reduce the incidence of contrast-related nephropathy, but this finding is reported inconsistently across currently available trials. High-quality, large clinical trials are needed before acetylcysteine use in this indication can be recommended universally."

Qualitative signal **Major change**: Definite benefit → possible benefit

Quantitative signal **Change in statistical significance**: the relative risk of contrast nephropathy with acetylcysteine versus hydration alone lost its statistical significance

RR of 0.44 (0.22, 0.88) → 0.81 (0.58, 1.13)

The loss of statistical significance first occurred with Gomes 2003, at which time the updated result was 0.61 (0.37, 1.00)

Change in effect magnitude of 50% or more: The relative risk reduction (RRR) decreased from 0.66 to 0.19

Other signals: Increase in number of trials of at least 50%: 7 trials → 17 trials (20 trials included in newer meta-analysis, but not all provided data on the primary outcome)

Increase in number of patients of at least 50%: N=805 → N=1964

Source(s) of Newer meta-analysis:
evidence Nallamothu BK, Shojania KG, Saint S, et al. Is acetylcysteine effective in preventing contrast-related nephropathy? A meta-analysis. *Am J Med* 2004;117:938-947.

This meta-analysis included 20 trials. The quantitative signal for change in statistical significance occurred with Gomes V, et al. Prevention of contrast-induced nephropathy with N-acetylcysteine in patients undergoing coronary angiography a randomized multicenter trial. Circulation 2003:108:IV–460.

Time to signal Qualitative signal: 1.4 years
Quantitative signal: -0.2 years

Original Review	Ducharme FM, Hicks GC. Anti-leukotriene agents compared to inhaled corticosteroids in the management of recurrent and/or chronic asthma (Cochrane Review). In: The Cochrane Library, Issue 1, 2001. Oxford: Update Software.
Question(s) addressed	How do anti-leukotriene agents compare with inhaled glucocorticoids in terms of efficacy and safety in the management of chronic asthma?
Findings of Original Review	The original review showed non significant trends towards superiority of inhaled corticosteroids, but found the evidence insufficient to permit reliable conclusions regarding relative efficacy of the two treatments. The reviewers concluded: "Anti-leukotriene agents had a similar rate of exacerbations compared to inhaled corticosteroids, but inhaled steroids produced better lung function and quality of life as well as reduced symptoms, night awakenings and need for rescue beta2-agonist. Reliable conclusions cannot yet be drawn regarding the efficacy of this treatment due to the paucity of trials published in full text."
New Findings	A subsequent update of the original review reported: "Patients treated with anti-leukotrienes were 60% more likely to suffer an exacerbation requiring systemic steroids…Significant differences favouring ICS were noted in most secondary outcomes, eg improvement in FEV1…symptom scores… Other significant benefits of ICS were seen for nocturnal awakenings, rescue medication use, and quality of life. Risk of side effects was not different between groups, but anti-leukotriene therapy was associated with 30% increased risk of "withdrawals for any cause" or "withdrawals due to poor asthma control". The updated review concluded, "For most asthma outcomes, ICS at 400 mcg/day of beclomethasone-equivalent are more effective than anti-leukotriene agents given in the usual licensed doses... Inhaled glucocorticoids should remain the first line monotherapy for persistent asthma."
Qualitative signal	**Major change:** possibly inferior → definitely inferior

Quantitative signal

Change in statistical significance: The risk of asthma exacerbations with anti-leukotrienes vs inhaled steroids (in adults and children) became statistically significant

Relative risk of 1.34 (0.93, 1.91) → 1.45 (1.07, 1.97)

Other signals: Increase in number of patients of at least 50%: N=1050 → N=1938

Increase in number of trials of at least 50%: 4 trials → 6 trials (for the above outcome)

Source(s) of Evidence

Subsequent meta-analysis (explicit update): Ducharme FM, Hicks GC. Anti-leukotriene agents compared to inhaled corticosteroids in the management of recurrent and/or chronic asthma in adults and children. *Cochrane Database Syst Rev*. 2002(3):CD002314.

New trials included in the meta-analysis

1. Bleecker ER, Welch MJ, Weinstein SF, et al. Low-dose inhaled fluticasone propionate versus oral zafirlukast in the treatment of persistent asthma. *J Allergy Clin Immunol*. 2000;105(6 Pt 1):1123-9.
2. Busse W, Raphael GD, Galant S, et al. Low-dose fluticasone propionate compared with montelukast for first-line treatment of persistent asthma: a randomized clinical trial. *J Allergy Clin Immunol*. 2001;107(3):461-8.
3. Kim KT, Ginchansky EJ, Friedman BF, et al. Fluticasone propionate versus zafirlukast: effect in patients previously receiving inhaled corticosteroid therapy. *Ann Allergy Asthma Immunol*. 2000;85(5):398-406.

Time to signal Qualitative signal: 2 years

Quantitative signal: 0.4 years (became positive with Bleecker 2000)

Example of criterion A5 **for 'Expansion of treatment':** Pivotal trial, new or discordant meta-analysis, trial indexed in *ACP J Club*, more recent practice guideline, or recent textbook has expanded of the role of the treatment (e.g., the treatment has now been shown to be of benefit in children or the elderly; or benefit now shown to apply to primary prevention of disease, not just secondary prevention).

Original Review

McIntyre PB, Berkey CS, King SM, et al. Dexamethasone as adjunctive therapy in bacterial meningitis. A meta-analysis of randomized clinical trials since 1988. JAMA 1997;278:925-931.

Question(s)

Does dexamethasone administered as an adjunct to antibiotic therapy improve

addressed	outcomes for patients with bacterial meningitis, and does effectiveness vary by subcategories of causative organisms and timing or nature of antibiotic therapy?
Findings of Original Review	The original review found that "in Haemophilus influenzae type b meningitis, dexamethasone reduced severe hearing loss overall (combined odds ratio [OR], 0.31; 95% confidence interval [CI], 0.14-0.69)" and "in pneumococcal meningitis, only studies in which dexamethasone was given early suggested protection, which was significant for severe hearing loss (combined OR, 0.09; 95% CI, 0.0-0.71) and approached significance for any neurological or hearing deficit (combined OR, 0.23; 95% CI, 0.04-1.05)." The authors concluded that "The available evidence on adjunctive dexamethasone therapy confirms benefit for H influenzae type b meningitis and, if commenced with or before parenteral antibiotics, suggests benefit for pneumococcal meningitis in childhood." The review contained only one study that included some adults (up to age 25 years of age).
New Findings	A pivotal trial that focused on adults patients and administered dexamethasone before or with the first dose of antibiotic and was given every 6 hours for four days showed: "treatment with dexamethasone was associated with a reduction in the risk of an unfavorable outcome (relative risk, 0.59; 95 percent confidence interval, 0.37 to 0.94; p =0.03). Treatment with dexamethasone was also associated with a reduction in mortality (relative risk of death, 0.48; 95 percent confidence interval, 0.24 to 0.96; p =0.04). Among the patients with pneumococcal meningitis, there were unfavorable outcomes in 26 percent of the dexamethasone group, as compared with 52 percent of the placebo group (relative risk, 0.50; 95 percent confidence interval, 0.30 to 0.83; p =0.006)." The authors concluded that "early treatment with dexamethasone improves the outcome in adults with acute bacterial meningitis and does not increase the risk of gastrointestinal bleeding."
Qualitative signal	**Major change**: **benefit reported in original review expanded to a new patient population** The original review concluded adjunctive dexamethasone conferred benefit only in children with acute bacterial meningitis due to Haemophilus influenzae type b and possibly pneumococcal meningitis. A pivotal trial showed significant benefit for adjunctive dexamethasone in adults with acute bacterial meningitis**.**
Quantitative signal	Not applicable
Other signals:	None
Source(s) of	**Pivotal trial:**

new evidence de Gans J, van de Beek D; European Dexamethasone in Adulthood Bacterial Meningitis Study Investigators. Dexamethasone in adults with bacterial meningitis. N Engl J Med. 2002;347(20):1549-56.

Time to signal Qualitative signal: 5.2 years
Quantitative signal: Not applicable

Example of criterion A6 for **Important caveat**: Pivotal trial, new meta-analysis, more recent practice guideline, or recent textbook adds an important caveat, about the patient populations who benefit, way in which treatment has to be delivered in order to derive benefit, sustainability of benefit (e.g., benefits on short term outcomes, but not long-term ones), or increases in harm that are not sufficient to undermine use altogether, but would clearly affect the decision to recommend treatment for at least some patient populations.

Original Review Abramson MJ, Puy RM, Weiner JM. Is allergen immunotherapy effective in asthma? A meta-analysis of randomized controlled trials. *Am J Respir Crit Care Med* 1995;151:969-974.

Question(s) addressed How efficacious is allergen immunotherapy in controlling the symptoms, improving lung function, or decreasing the requirements for medication use in patients with asthma?

Findings of Original Review The original review included 20 randomized placebo controlled double-blind trials and reported that "combined odds of symptomatic improvement from immunotherapy with any allergen were 3.2 (95% CI 2.2 to 4.9). The odds for reduction in medication after mite immunotherapy were 4.2 (95% CI 2.2 to 7.9). The combined odds for reduction in BHR [bronchial hyperreactivity] were 6.8 (95% CI 3.8 to 12.0). The mean effect size for any allergen immunotherapy on all continuous outcomes was 0.71 (95% CI 0.43 to 1.00), which would correspond to a mean 7.1% predicted improvement in FEV1 from immunotherapy."

The authors also pointed out that "Although the benefits of allergen immunotherapy could be overestimated because of unpublished negative studies, an additional 33 such studies would be necessary to overturn these results." They thus concluded that "allergen immunotherapy is a treatment option in highly selected patients with extrinsic ("allergic") asthma."

New Findings A pivotal trial reported that: "During the two treatment years, the mean peak expiratory flow rate was higher in the immunotherapy group (489 +/- 16 liters per minute, vs. 453 +/- 17 in the placebo group [p =0.06] during the first year, and 480 +/- 12 liters per minute, vs. 461 +/- 13 in the placebo group [p =0.03] during the second). Medication use was higher in the

immunotherapy group than in the placebo group during observation and lower during the first treatment year (p =0.01) but did not differ in the two groups during the second year (p =0.7). Asthma-symptom scores were similar in the two groups (p =0.08 in year 1 and p =0.3 in year 2). The immunotherapy group had reduced hay-fever symptoms, skin-test sensitivity to ragweed, and sensitivity to bronchial challenges and increased IgG antibodies to ragweed as compared with the placebo group; there was no longer a seasonal increase in IgE antibodies to ragweed allergen in the immunotherapy group after two years of treatment. Reduced medication costs were counterbalanced by the costs of immunotherapy."

The authors concluded that "Although immunotherapy for adults with asthma exacerbated by seasonal ragweed exposure had positive effects on objective measures of asthma and allergy, the clinical effects were limited and many were not sustained for two years."

Qualitative signal **Major change**: **important caveat**
In this case, the caveat concerns the sustainability of benefit.

Quantitative signal None met

Other signals: None

Source(s) of new evidence Pivotal trial: Creticos PS, Reed CE, et al. Ragweed immunotherapy in adult asthma. N Engl J Med. 1996;334(8):501-6.

Time to signal Qualitative signal: 327 days
Quantitative signal: Not applicable

Example of criterion A7 for Opposing findings from discordant meta-analysis or non-pivotal trial: The treatment has been characterized in sufficiently different terms to the cohort review that disagreement would have met criteria for 'potentially invalidating change' (A1) except the source was not a pivotal trial, new-meta-analysis, or more recent practice guideline, or recent textbook—rather, it was a discordant meta-analysis or trial indexed in *ACP J Club*.

Original Review Hood SC, Moher D, Barber GG. Management of intermittent claudication with pentoxifylline: meta-analysis of randomized controlled trials. *CMAJ*. 1996;155(8):1053-9.

Question(s) addressed Does pentoxifylline improve the walking capacity of patients with moderate intermittent claudication?

Findings of The original meta-analysis found " a statistically significant improvement in

Original Review the pain-free walking distance after pentoxifylline therapy (weighted mean difference 29.4 m [95% confidence interval (CI) 13.0 to 45.9 m])… A significant improvement was also noted in the absolute claudication distance (weighted mean difference 48.4 m [95% CI 18.3 to 78.6 m])". The authors concluded that "pentoxifylline therapy may be efficacious in improving the walking capacity of patients with moderate intermittent claudication."

New Findings A randomized trial with a commentary in *ACP Journal Club* (Dawson 2002) compared pentoxifylline with an alternative medication, cilostazol, and placebo. The authors reported: "Mean maximal walking distance of cilostazol-treated patients (n = 227) was significantly greater at every postbaseline visit compared with patients who received pentoxifylline (n = 232) or placebo (n = 239). After 24 weeks of treatment, mean maximal walking distance increased by a mean of 107 m (a mean percent increase of 54% from baseline) in the cilostazol group, significantly more than the 64-m improvement (a 30% mean percent increase) with pentoxifylline ($p < 0.001$). The improvement with pentoxifylline was similar ($p = 0.82$) to that in the placebo group (65 m, a 34% mean percent increase)."

The authors concluded that "Cilostazol was significantly better than pentoxifylline or placebo for increasing walking distances in patients with intermittent claudication… Pentoxifylline and placebo had similar effects."

The seventh ACCP Conference on Antithrombotic and Thrombolytic Therapy (Clagett 2004) and *UpToDate* characterize pentoxifylline as no better than exercise and quote the above trial as the basis for this assessment.

Qualitative signal **Major change**: possibly beneficial → definitely not beneficial

The original review concluded that pentoxifylline was likely efficacious in the treatment of intermittent claudication. A major practice guideline and chapter in recent textbook characterize pentoxifylline as no better than placebo based on the results of a trial that did not meet criteria for pivotal but was indexed in *ACP Journal Club.*

Quantitative signal None met

Other signals: None

Source(s) of new evidence **Trial indexed in ACP J Club:**

Dawson DL, Cutler BS, Hiatt WR, et al. A comparison of cilostazol and pentoxifylline for treating intermittent claudication. Am J Med. 2000;109(7):523-30.

Practice guideline:

Clagett GP, Sobel M, Jackson MR, Lip GY, Tangelder M, Verhaeghe R. Antithrombotic therapy in peripheral arterial occlusive disease: the Seventh

ACCP Conference on Antithrombotic and Thrombolytic Therapy. Chest. 2004;126(3 Suppl):609S-626S.

Time to signal Qualitative signal: 4.1 years
Quantitative signal: not applicable

Appendix C: Sample Subject Searches

Examples of subject searches and how search features are scored. The last line of each search retrieved the Clinical Query set. The second from last line retrieves the Core Clinical Journal RCT set. The third from last line retrieves the meta-analysis set.

Eikelboom JW Quinlan DJ and Douketis JD, Extended-duration prophylaxis against venous thromboembolism after total hip or knee replacement: a meta-analysis of the randomised trials. Lancet. 2001 Jul 7;358(9275):9-15.

Feature scoring: 5 terms and exploded terms (lines 1,2, 3). The clinical query used is the sensitivity query.

1. exp Heparin/

2. exp Warfarin/

3. exp Venous Thrombosis/

4. Arthroplasty, Replacement, Knee/

5. Arthroplasty, Replacement, Hip/

6. 1 or 2

7. 4 or 5

8. 3 and 6 and 7

9. limit 8 to (yr="2000 - 2006" and meta analysis)

10. limit 8 to ("core clinical journals (aim)" and yr="2000 - 2006" and randomized controlled trial)

11. limit 8 to ("therapy (specificity)" and yr="2000 - 2006")

Fowlie PW., Prophylactic indomethacin: systematic review and meta-analysis, Arch Dis Child Fetal Neonatal Ed. 1996 Mar;74(2):F81-7.

Feature scoring: 4 terms, exploded terms (line 1) and free text terms (line 3)

1. exp Indomethacin/

2. Ductus Arteriosus, Patent/

3. Cerebral Hemorrhage/ or intraventricular hemorrhage.mp.

4. 2 or 3

5. 1 and 4

6. limit 5 to (yr="1994 - 2006" and meta analysis)

7. limit 5 to ("core clinical journals (aim)" and yr="1994 - 2006" and randomized controlled trial)

8. limit 5 to ("therapy (optimized)" and yr="1994 - 2006")

Bucher HC, Hengstler, P, Schindler C, and Guyatt GH. Percutaneous transluminal coronary angioplasty versus medical treatment for non-acute coronary heart disease: meta-analysis of randomised controlled trials, BMJ. 2000 Jul 8;321(7253):73-7.

Feature scoring: 2 terms, starred terms (lines 1 and 2) and subheadings (line 2).

1. *Angioplasty, Transluminal, Percutaneous Coronary/

2. *Myocardial Infarction/tu, dt

3. 1 and 2

4. limit 3 to (yr="1995 - 2006" and meta analysis)

5. limit 3 to ("core clinical journals (aim)" and yr="1995 - 2006" and randomized controlled trial)

6. limit 3 to ("therapy (optimized)" and yr="1995 - 2006")

Appendix D: Sample Assessment Worksheet

From Cohort ID 167

Fowlie PW., Prophylactic indomethacin: systematic review and meta-analysis, Arch Dis Child Fetal Neonatal Ed. 1996 Mar;74(2):F81-7.

Cohort Refid	Study refid	Author	Pub Year	MHDA	PMID	Type	On topic?	Eligible	N	Notes
167	1	Mahony	1985	01/05/1985	3998921	Original			104	
167	2	Ment	1985	01/12/1985	3906073	Original			48	N for original is taken from Figure 2 unless otherwise noted
167	3	Vincer	1987	01/11/1987	3321891	Original			30	
167	4	Rennie	1986	01/03/1986	3516077	Original			40	
167	5	Krueger	1987	01/11/1987	3312552	Original			32	
167	6	Hanigan	1988	01/06/1988	3373404	Original			112	
167	7	Ment	1988	01/06/1988	3373405	Original			36	
167	8	Bandstra	1988	01/10/1988	3174314	Original			199	
167	9	Bada	1989	01/10/1989	2677294	Original			141	
167	10	Ment	1994	01/06/1994	8201485	Original			61	
167	11	Ment	1994	01/04/1994	8134206	Original			431	
167	12	Setzer	1984		Not in MEDLINE	Original			59	from table 2 of refid 167
167	13	Puckett	1985		Not in MEDLINE	Original			32	
167	14	Bandstra	1987		Not in MEDLINE	Original			199	from table 2
167	76	Ohlsson	2005	24/02/2006	16235321	MA	N			
167	77	Loe	2005	11/08/2005	15994634	MA	N			
167	78	Thomas	2005	21/07/2005	15717178	MA	Y	N		
167	79	Simmer	2005	19/07/2005	15846747	MA	N			
167	80	Stevens	2004	30/11/2004	15266470	MA	N			
167	81	Herrera	2004	29/06/2004	14974018	MA	Y	N		New comparision b/w prolonged and short course of Indomethacin without definitive results
167	82	Ohlsson	2004	29/06/2004	14973955	MA	N			
167	83	Fowlie	2003	10/12/2003	14602691	MA	Y	Y		
167	84	Cooke	2003	28/07/2003	12804488	MA	Y	N		no new study
167	85	Shah	2003	27/03/2003	12535425	MA	N			
167	86	Brion	1999	26/07/1999	10353408	MA	Y	N		
167	87	Rubino	1997	05/06/1997	9165939	MA	Y	N		
167	88	Clark	1996	03/01/1997	8951253	MA	N			
167	89	Clyman	1996	26/06/1996	8627430	MA	Y			
167	74	Simko	1994	01/06/1994	8169675	Candidate				
167	72	Rush	1994	29/09/1994	8071758	Candidate				
167	71	Hammerman	1995	27/02/1995	7838642	Candidate				
167	70	Ment	1995	14/04/1995	7892866	Candidate				
167	69	Corbet	1995	07/07/1995	7776110	Candidate				
167	68	Malloy	1995	11/10/1995	7666265	Candidate				
167	66	Bernstein	1996	07/06/1996	8618177	Candidate				
167	67	Van	1995	07/06/1996	8618789	Candidate				
167	65	Couser	1996	26/06/1996	8627434	Candidate	Y	Y	90	Identified in ref id 83 and included in its

Cohort Refid	Study refid	Author	Pub Year	MHDA	PMID	Type	On topic?	Eligible	N	Notes
										pooled stats
167	64	Ment	1996	01/07/1996	8632956	Candidate				
167	63	Silver	1996	15/08/1996	8677067	Candidate				
167	62	Rastogi	1996	29/08/1996	8692619	Candidate				
167	61	Ment	1996	04/12/1996	8885951	Candidate				
167	60	Gerstmann	1996	03/01/1997	8951252	Candidate				
167	59	Yaseen	1997	12/05/1997	9078828	Candidate	Y	Y	27	Identified in ref id 83 and included in its pooled stats
167	58	Van	1997	26/06/1997	9175948	Candidate				
167	57	Allan	1997	10/07/1997	9193243	Candidate				
167	56	Hudak	1997	21/07/1997	9200358	Candidate				
167	55	Parilla	1997	04/08/1997	9215169	Candidate				
167	54	Romagnoli	1997	11/09/1997	9284854	Candidate				
167	53	Keszler	1997	14/10/1997	9310511	Candidate				
167	52	Chugh	1997	13/11/1997	9332109	Candidate				
167	50	Su	1998	30/04/1998	9553290	Candidate				
167	49	Morales	1998	21/05/1998	9568218	Candidate				
167	48	Ment	1998	14/08/1998	9714645	Candidate				
167	51	Yeh	1997	10/09/1998	9310536	Candidate				
167	47	rroyo-Cabrales	1998	12/11/1998	9775459	Candidate				
167	46	Anand	1999	16/04/1999	10201714	Candidate				
167	45	Tammela	1999	25/05/1999	10228288	Candidate				
167	44	Kothadia	1999	22/07/1999	10390255	Candidate				
167	43	Vohr	1999	26/08/1999	10405190	Candidate				
167	42	Panter	1999	02/09/1999	10430197	Candidate				
167	41	Kopelman	1999	12/10/1999	10484801	Candidate				
167	40	Baenziger	1999	11/01/2000	10592922	Candidate				
167	36	Patel	2000	20/01/2000	10625080	Candidate				
167	35	Ment	2000	22/03/2000	10699097	Candidate				
167	39	Su	1999	23/03/2000	10525023	Candidate				
167	38	Supapannachart	1999	05/04/2000	10730525	Candidate	Y	Y	30	Identified in ref id 83 and included in its pooled stats
167	34	Couser	2000	21/06/2000	10850507	Candidate	Y	N		Outcome not considered
167	33	Van	2000	07/09/2000	10974130	Candidate				
167	32	De Carolis	2000	14/09/2000	10834523	Candidate				
167	37	Sanghvi	1999	05/01/2001	10740301	Candidate				
167	30	Stark	2001	11/01/2001	11150359	Candidate				
167	29	Wardle	2001	22/02/2001	11124916	Candidate				
167	28	Van	2001	05/04/2001	11174617	Candidate				
167	31	Dani	2000	31/05/2001	11106052	Candidate				
167	27	Schmidt	2001	05/07/2001	11430325	Candidate	Y	Y	1202	Identified in ref id 83 and included in its pooled stats
167	26	Vermont	2001	18/10/2001	11533345	Candidate				
167	25	Schmidt	2002	07/10/2002	12241754	Candidate				
167	24	Christmann	2002	22/11/2002	12061361	Candidate				
167	23	Lago	2002	04/02/2003	12014386	Candidate	Y	N		
167	22	Supapannachart	2002	12/02/2003	12549803	Candidate	Y	N		
167	21	Vohr	2003	18/04/2003	12671149	Candidate	Y	N		

Cohort Refid	Study refid	Author	Pub Year	MHDA	PMID	Type	On topic?	Eligible	N	Notes
167	20	Lee	2003	03/10/2003	12897285	Candidate	Y	N		
167	19	Simons	2003	25/11/2003	14612478	Candidate	N			
167	18	Schreiber	2003	08/12/2003	14645637	Candidate	N			
167	17	Osborn	2003	10/12/2003	14602694	Candidate	Y	N		
167	16	Chotigeat	2003	05/02/2004	14700149	Candidate	Y	N		
167	15	Su	2003	26/02/2004	14651538	Candidate	Y	N		
167	14	Kumar	2004	05/10/2004	15235161	Candidate	Y	Y	115	
167	13	Gournay	2004	15/12/2004	15567009	Candidate	N			
167	12	Van	2004	15/12/2004	15567010	Candidate	N			
167	11	Vila-Vazqu	2004	18/03/2005	15298714	Candidate	N			
167	10	Mestan	2005	12/07/2005	16000353	Candidate	N			
167	9	Van Meurs	2005	12/07/2005	16000352	Candidate	N			
167	8	Hall	2005	27/09/2005	15867047	Candidate	N			
167	6	Bell	2005	10/11/2005	15930233	Candidate	N			
167	7	Dani	2005	10/11/2005	15930213	Candidate	N			
167	5	Gimeno	2005	07/12/2005	16219273	Candidate	Y	N		
167	4	Olney	2005	29/12/2005	16041635	Candidate	N			
167	3	Zanardo	2005	19/01/2006	16260891	Candidate	Y	N		
167	1	Mercer	2006	10/05/2006	16585320	Candidate	N			
167	2	Adamska	2005	25/05/2006	16547381	Candidate	Y	N		
167		Morales-Suarez	1994		Not in MEDLINE	Nominatio n	Y	Y	80	Identified in Cochrane version of Fowlie 2003. Analysis 01.02
167		Domanico	1994		Not in MEDLINE	Nominatio n	Y	Y	100	Fowlie 2003 Analysis 01.04

Note candidate studies after the last search date of ref id 83 were screened.

Appendix E: Searches by Clinical Area

Searches for each clinical area (a complete example, showing publication type and year limits is shown below.

Cardiac and cardiovascular systems

1. exp Heart Diseases/
2. Heart Failure, Congestive/
3. or/1-2

Neurology

1. exp Nervous System Diseases/

Gastroenterology and Hepatology

1. exp Gastrointestinal Diseases/

Critical Care Medicine

1. exp Critical Care/
2. exp Critical Illness/

Endocrinology and Metabolism

1. exp Bone Diseases, Metabolic/
2. exp Metabolic Diseases/
3. or/1-2

Infectious Diseases

1. exp Anti-infective agents/

Obstetrics and Gynecology

1. exp Genital Diseases, Female/
2. exp Pregnancy Complications/
3. or/1-2

Oncology

1. exp Neoplasms/

Peripheral Vascular Diseases

1. exp Vascular Diseases/

Psychiatry

1. exp Mental Disorders/

Respiratory System

1. exp Respiratory Tract Diseases/

Rheumatology

1. exp Rheumatic Diseases/

Urology and Nephrology

1. exp Kidney Diseases/

Example of a complete search:
Cardiac and cardiovascular systems

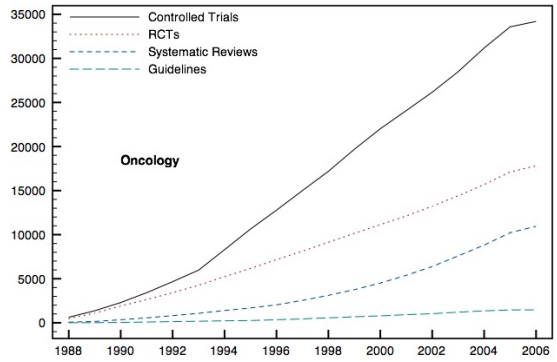
1. exp Heart Diseases/
2. Heart Failure, Congestive/
3. or/1-2
4. limit 3 to animals
5. remove duplicates from 4
6. limit 5 to yr="1988"
7. limit 5 to yr="1989"
8. limit 5 to yr="1990"
9. limit 5 to yr="1991"
10. limit 5 to yr="1992"
11. limit 5 to yr="1993"
12. limit 5 to yr="1994"
13. limit 5 to yr="1995"
14. limit 5 to yr="1996"
15. limit 5 to yr="1997"
16. limit 5 to yr="1998"
17. limit 5 to yr="1999"
18. limit 5 to yr="2000"
19. limit 5 to yr="2001"
20. limit 5 to yr="2002"
21. limit 5 to yr="2003"
22. limit 5 to yr="2004"
23. limit 5 to yr="2005"
24. limit 5 to yr="2006"
25. limit 5 to yr="2007"

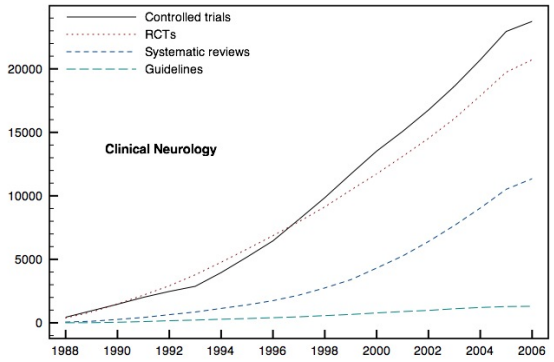
26. limit 3 to randomized controlled trial
27. limit 3 to clinical trial
28. 27 not 26
29. remove duplicates from 28
30. limit 29 to yr="1988"
31. limit 29 to yr="1989"
32. limit 29 to yr="1990"
33. limit 29 to yr="1991"
34. limit 29 to yr="1992"
35. limit 29 to yr="1993"
36. limit 29 to yr="1994"
37. limit 29 to yr="1995"
38. limit 29 to yr="1996"
39. limit 29 to yr="1997"
40. limit 29 to yr="1998"
41. limit 29 to yr="1999"
42. limit 29 to yr="2000"
43. limit 29 to yr="2001"
44. limit 29 to yr="2002"
45. limit 29 to yr="2003"
46. limit 29 to yr="2004"
47. limit 29 to yr="2005"
48. limit 29 to yr="2006"
49. limit 29 to yr="2007"
50. limit 3 to randomized controlled trial
51. remove duplicates from 50
52. limit 51 to yr="1988"
53. limit 51 to yr="1989"
54. limit 51 to yr="1990"
55. limit 51 to yr="1991"
56. limit 51 to yr="1992"
57. limit 51 to yr="1993"
58. limit 51 to yr="1994"
59. limit 51 to yr="1995"
60. limit 51 to yr="1996"
61. limit 51 to yr="1997"
62. limit 51 to yr="1998"
63. limit 51 to yr="1999"
64. limit 51 to yr="2000"
65. limit 51 to yr="2001"
66. limit 51 to yr="2002"
67. limit 51 to yr="2003"
68. limit 51 to yr="2004"
69. limit 51 to yr="2005"
70. limit 51 to yr="2006"
71. limit 51 to yr="2007"
72. limit 3 to systematic reviews
73. remove duplicates from 72
74. limit 73 to yr="1988"
75. limit 73 to yr="1989"
76. limit 73 to yr="1990"
77. limit 73 to yr="1991"
78. limit 73 to yr="1992"
79. limit 73 to yr="1993"
80. limit 73 to yr="1994"
81. limit 73 to yr="1995"
82. limit 73 to yr="1996"
83. limit 73 to yr="1997"
84. limit 73 to yr="1998"
85. limit 73 to yr="1999"
86. limit 73 to yr="2000"
87. limit 73 to yr="2001"
88. limit 73 to yr="2002"
89. limit 73 to yr="2003"
90. limit 73 to yr="2004"
91. limit 73 to yr="2005"
92. limit 73 to yr="2006"
93. limit 73 to yr="2007"
94. limit 3 to guideline
95. remove duplicates from 94
96. limit 95 to yr="1988"
97. limit 95 to yr="1989"
98. limit 95 to yr="1990"
99. limit 95 to yr="1991"
100. limit 95 to yr="1992"
101. limit 95 to yr="1993"
102. limit 95 to yr="1994"
103. limit 95 to yr="1995"
104. limit 95 to yr="1996"
105. limit 95 to yr="1997"
106. limit 95 to yr="1998"
107. limit 95 to yr="1999"
108. limit 95 to yr="2000"
109. limit 95 to yr="2001"
110. limit 95 to yr="2002"
111. limit 95 to yr="2003"
112. limit 95 to yr="2004"
113. limit 95 to yr="2005"
114. limit 95 to yr="2006"
115. limit 95 to yr="2007"

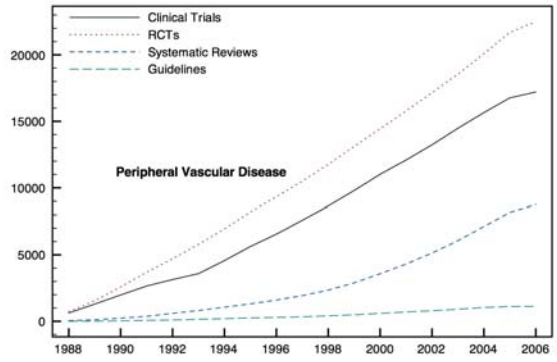
Appendix F: Growth of the Literature by Clinical Area

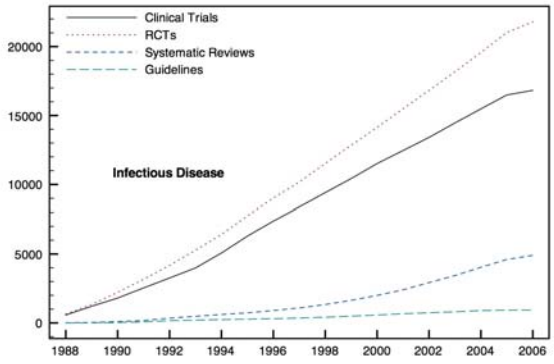
Cumulative publications by type and clinical area

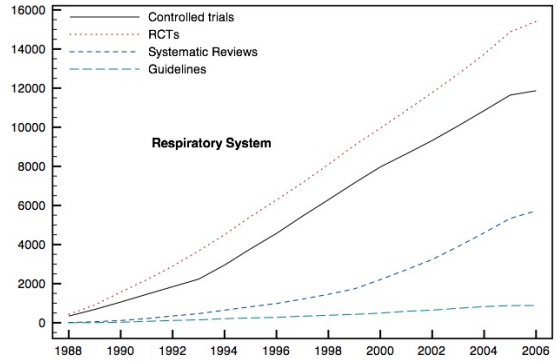
X axis represents number of records, Y axis is year in each case.

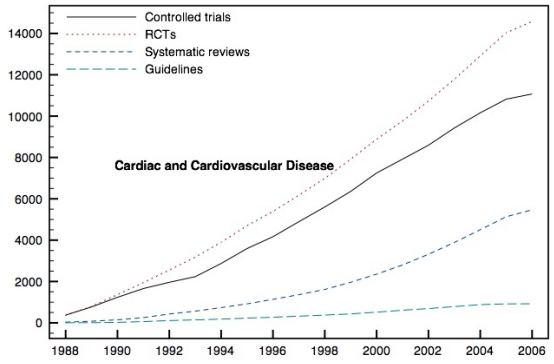


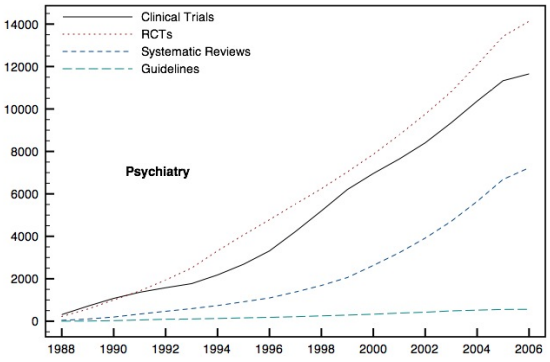


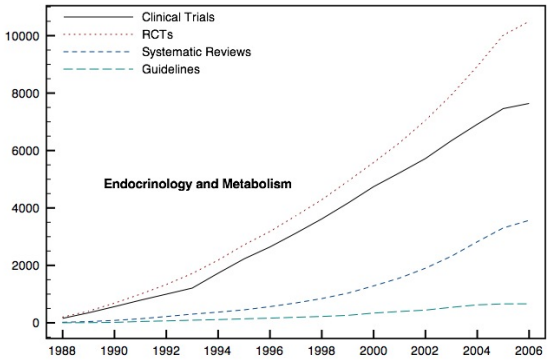


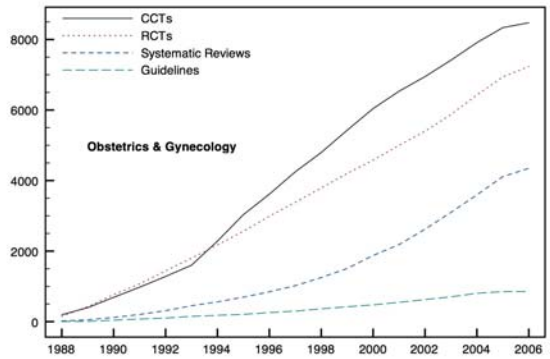


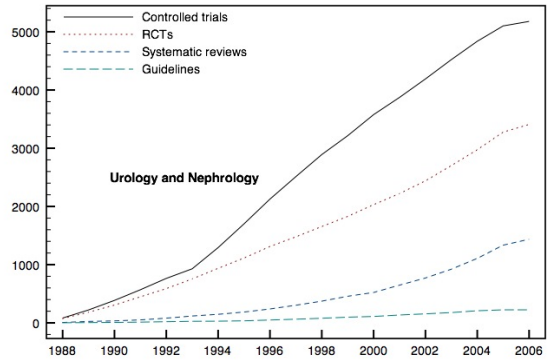


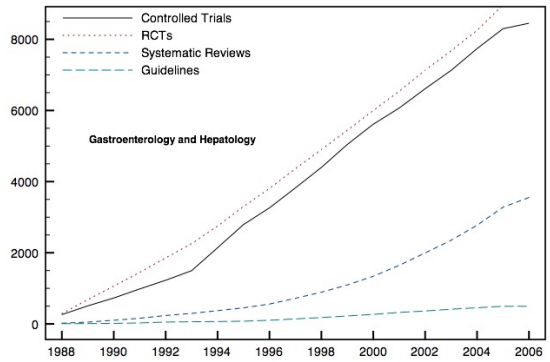


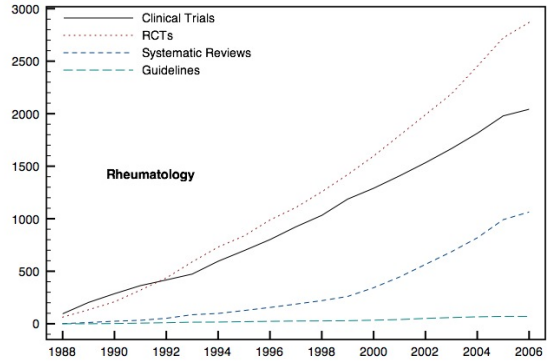

Clinical Trials
RCTs
Systematic Reviews
Guidelines
Psychiatry
14000
12000
10000
8000
6000
4000
2000
0
1988
1990
1992
1994
1996
1998
2000
2002
2004
2006

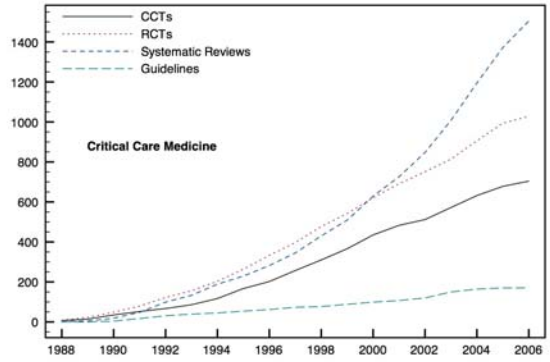

Clinical Trials
RCTs
Systematic Reviews
Guidelines
Endocrinology and Metabolism
10000
8000
6000
4000
2000
0
1988
1990
1992
1994
1996
1998
2000
2002
2004
2006


CCTs
RCTs
Systematic Reviews
Guidelines
Obstetrics & Gynecology
8000
6000
4000
2000
0
1988
1990
1992
1994
1996
1998
2000
2002
2004
2006


Controlled trials
RCTs
Systematic reviews
Guidelines
Urology and Nephrology
5000
4000
3000
2000
1000
0
1988
1990
1992
1994
1996
1998
2000
2002
2004
2006


Controlled Trials
RCTs
Systematic Reviews
Guidelines
Gastroenterology and Hepatology
8000
6000
4000
2000
0
1988
1990
1992
1994
1996
1998
2000
2002
2004
2006


Clinical Trials
RCTs
Systematic Reviews
Guidelines
Rheumatology
3000
2500
2000
1500
1000
500
0
1988
1990
1992
1994
1996
1998
2000
2002
2004
2006


CCTs
RCTs
Systematic Reviews
Guidelines
Critical Care Medicine
1400
1200
1000
800
600
400
200
0
1988
1990
1992
1994
1996
1998
2000
2002
2004
2006

Appendix G: Survey Instrument

Wednesday, March 14, 2007

Updating Systematic Reviews Survey (University of Ottawa EPC)

Part I: Survey Information & Consent Form

Survey Title: Examining the Updating Practices of Healthcare Organizations that Fund and/or Conduct Systematic Review Research.

Background: To maintain their central role in informing clinical practice and health care policy, systematic reviews need to be up-to date. However, there is little available guidance about the appropriate timing or approach for updating of a particular evidence-base. Furthermore, updating policies and practices of organizations are often not explicit. There are general concerns about dated reviews; however, unlike other methodological areas, methods for updating systematic reviews remain understudied. Invitation to Participate: Your organization/group/unit has been chosen to participate in the first phase of this international survey on systematic review updating experiences given its involvement as a producer and/or funder of individual systematic reviews, or systematic reviews that are incorporated into clinical practice guidelines (CPGs) or health technology assessments (HTAs). Benefits: To better understand the updating experiences of those engaged in systematic reviews research, your responses would facilitate the advancement of best practice of updating methodology. This research is also of relevance to evidence-based agencies and researchers wishing to increase general performance and efficiency in their respective updating practices.

This research is funded by the U.S. Agency for Healthcare Research and Quality (AHRQ) Evidence-based Practice Center (EPC) Program and is being undertaken by the University of Ottawa Evidence-based Practice Center (UO-EPC) as part of an ongoing research program to examine a broad spectrum of issues pertaining to the methodology for updating systematic reviews. This will also be part of my MSc thesis in Public Health Sciences, University of Toronto, Canada. The Principle Investigator of this study is Dr. David Moher, Director of the UO-EPC and the Chalmers Research Group, based at the Children's Hospital of Eastern Ontario Research Institute (CHEO RI).

Instructions: As a participant, we are asking that you complete this survey, which is anticipated to take 30 minutes. Your organization's participation is voluntary and at any time you can leave the survey by clicking on "Exit this survey". If you wish to return, your responses will be saved. You may withdraw from the study at any time and may refrain from answering any question by leaving it blank. All responses will remain strictly confidential and we anticipate minimal or no risk of harm from the research study. As a small token of our appreciation for completing this survey, you will be provided with a $10 gift certificate from Amazon.com. In order to obtain your consent, please review each statement below:

1. I am over 18 years of age and have the approval from my organization/group/unit to represent them as a participant of this survey;

2. I understand that the information obtained from the study will be used within a MSc thesis and that aggregate results from this survey may be published in peer-reviewed journals, however will not be directly identify organizations or groups by name in any reports or publications;
3. I understand that my organization's participation is voluntary and withdrawal from the study is allowed at any stage without adverse consequences for my organization;
4. I understand in representing my organization, that I may refrain from answering any questions should I so wish;
5. I understand that apart from the $10 Amazon gift card being offered for completing this survey, that I may not directly benefit from taking part in the study and that I will not receive any payment for participating in this research;
6. I understand that my organization will be provided with an executive summary of the results upon request;
7. I understand the Children's Hospital of Eastern Ontario and the University of Toronto's Research Ethics Boards have formally approved this survey;
8. I understand that data collected for this research will be retained for a period of 3 years, and will be stored in a secure location;
9. I understand that all responses will be kept strictly confidential;
10. I understand that I waive no legal rights by participating in this survey.

If you have any questions related to this survey, please contact:

Chantelle Garritty
University of Ottawa EPC Coordinator
CHEO RI, Ottawa, Canada
Email:cgarritty@cheo.on.ca
Tel: 613-737-7600 x 4117
Fax: 613-738-4800

If you have any questions or concerns about your rights as a research participant, please contact Sharon Haig (shaig@cheo.on.ca).

On behalf of the UO-EPC, thank you in advance for your time.
* I AGREE TO TAKE PART IN THIS STUDY (Confirming my consent to participate)

Yes _____ No _____

On behalf of the University of Ottawa EPC, thank you for your response.

If you still wish to exit this survey, please click on the Exit This Survey link provided above in the upper right corner of the screen.

Otherwise, click on 'PREVIOUS' to return to the preceding page to agree to take part in this survey.

Please review the following before starting the survey.

Systematic Review refers to a study that has a clear and comprehensive methodology. It includes explicit research questions, methods of searching the literature, criteria for including material, criteria for appraising quality and reliability of studies, and synthesis of research findings.

Please Note: For the purposes of this survey, the term Systematic Review is intended to be used broadly to reflect the following:

i) Conduct of individual or stand alone systematic reviews or
ii) Conduct of systematic reviews for use in Clinical Practice Guidelines (CPGs), or Health Technology Assessments (HTAs)

Therefore, please complete this survey based on the overall research synthesis work with which your organization/group/unit may be involved either as a producer, funder and/or both.

1. Which of the following statements best describe your organization/group? (Please tick one of the following options)

- ☐ A funder of systematic reviews (including as part of Clinical Practice Guidelines (CPGs), or Health Technology Assessments (HTAs))
- ☐ A producer of systematic reviews (including as part of Clinical Practice Guidelines (CPGs), or Health Technology Assessments (HTAs))
- ☐ Both a funder & producer of systematic reviews (including as part of Clinical Practice Guidelines (CPGs), or Health Technology Assessments (HTAs))

2. Which of the following statements best describe why your organization/group produces systematic reviews? (Please tick one of the following options)

- ☐ For knowledge support (e.g., Cochrane reviews or other discrete systematic reviews)
- ☐ For decision support (e.g., policy or clinical decision-making; evidence for HTA's or CPG's)
- ☐ For both knowledge & decision support
- ☐ Other (please specify)

3. We suggest UPDATE refers to "a discrete event aiming to search for and identify new evidence to incorporate into a previously completed systematic review." (New evidence is taken to mean any evidence not included in the previously completed review, irrespective of its' chronological appearance in the literature)

To what extent do you think your organization/group would agree or disagree with this definition? (Please tick one of the following options):

- ☐ Strongly agree
- ☐ Somewhat agree
- ☐ Neither agree nor disagree
- ☐ Somewhat disagree
- ☐ Strongly disagree
- ☐ No opinion

4. Overall, what level of importance would you estimate your organization/group places on updating systematic reviews? (Please tick one of the following options)

- ☐ 1 (low) 2 3 4 5 (high)

5. How would you rank updating of systematic reviews compared to other research mandates within your organization/group? (Please tick one of the following options)

□ 1 (low) 2 3 4 5 (high)

6. Does your organization/group have a policy on updating systematic reviews?

□ Yes
□ No
□ Not sure

7. Is establishing a formal updating policy or process something your organization/group would view as important? (Please tick one of the following options)

□ Yes
□ No
□ Not sure
□ Not applicable

8. Please describe the updating policy of your organization/group (if applicable):
(or list any working documents detailing the updating policy of your organization in the box provided below)

9. In general, how would you describe the current 'updating' practices of your organization/group? (Please tick one of the following options)

□ Update regularly
□ Update irregularly
□ Do not update

10. Of the systematic reviews either produced or commissioned by your organization/group, what percentage would you estimate are up to date?

□ 0-10%
□ 11-20%
□ 21-30%
□ 31-40%
□ 41-50%
□ 51-60%
□ 61-70%
□ 71-80%
□ 81-90%
□ 91-100%
□ Not sure

11. Which of the following does your organization/group feel is the most responsible for ensuring systematic reviews are updated?

□ Funder(s) of original review
□ Authors of the original review
□ Information specialist/Librarian
□ Policy-maker utilizing the evidence
□ All of the above

- None of the above
- Other (please specify)

12. How would you describe the method(s) used by your organization/group to determine the need to update a systematic review?

- Formal method(s) used
- Informal method(s) used
- No method(s) used
- Not sure

13. Does your organization/group conduct literature searches regularly to help identify new relevant literature?

- Yes
- No
- Not sure
- Not applicable

14. How often are searches conducted to identify new literature?
Please tick one of the following options

- Monthly
- Every 6 months
- Every 12 months
- Every 18 months
- Not applicable
- Other time interval (please specify)

15. How often are the following search strategies typically used when monitoring the literature to identify new evidence? Please tick one answer for each type of strategy

	Always	Often	Sometimes	Seldom	Never	Not sure	Not applicable
- Same search strategy from the original review	☐	☐	☐	☐	☐	☐	☐
- Modified search strategy from the original	☐	☐	☐	☐	☐	☐	☐
- New search strategy - original search	☐	☐	☐	☐	☐	☐	☐

16. How frequent does your organization/group use the following strategies to monitor emergence of new evidence for updating? (Please provide one answer for each strategy.)

	Always	Often	Sometimes	Seldom	Never	Not sure	Not applicable
- General literature searches (e.g. electronic, hand-search etc.	☐	☐	☐	☐	☐	☐	☐
- Trial registry surveillance (e.g., ClinicalTrials.Gov, WHO International Clinical Trials Registry, various national	☐	☐	☐	☐	☐	☐	☐

government or industry trials registries)							
- ***Statistical approaches*** (e.g., Barrowman's diagnostic test, cumulative meta-analysis etc.)	☐	☐	☐	☐	☐	☐	☐
- ***Systematic reviews*** surveillance	☐	☐	☐	☐	☐	☐	☐
Guideline or health technology assessment surveillance	☐	☐	☐	☐	☐	☐	☐
- Experts in the field	☐	☐	☐	☐	☐	☐	☐

17. Please list any additional strategies, actions, steps or methods used by your organization/group to monitor for new evidence,
(or list any working documents your organization has developed listing these monitoring strategies in the box provide below). If there are no additional strategies, actions, or steps that you are aware of, please leave blank.

18. How often do the following issues factor into your organization/group determining WHEN to conduct or fund an update? Please provide one answer for each reason

	Always	Often	Sometimes	Seldom	Never	Not sure	Not applicable
- Number of new studies identified	☐	☐	☐	☐	☐	☐	☐
- Number of participants in new studies	☐	☐	☐	☐	☐	☐	☐
- Reporting of serious or 'new' serious adverse events	☐	☐	☐	☐	☐	☐	☐
- New inclusion criteria (outcomes; interventions; populations; methodological advances/new analysis)	☐	☐	☐	☐	☐	☐	☐
- Totality (comprehensiveness) of all new evidence or data (including harms & Benefits)	☐	☐	☐	☐	☐	☐	☐
- Time credibility	☐	☐	☐	☐	☐	☐	☐
- Need for an internal organizational decision	☐	☐	☐	☐	☐	☐	☐
- Formal request from policy-maker or healthcare decision-maker	☐	☐	☐	☐	☐	☐	☐

19. How often do each of the following individuals/groups impact your organization/ group's decision-making process of whether to fund or conduct an update?
Please provide one answer for each reason

	Always	Often	Sometimes	Seldom	Never	Not sure	Not applicable
- Original authors of the systematic review (in-house or external to your organization)	☐	☐	☐	☐	☐	☐	☐
- Information specialist/Librarian	☐	☐	☐	☐	☐	☐	☐
- Experts in the field	☐	☐	☐	☐	☐	☐	☐
- Statistician(s)	☐	☐	☐	☐	☐	☐	☐
- External Policy-maker(s)	☐	☐	☐	☐	☐	☐	☐

- The organization/group itself as original funder of a systematic review	☐	☐	☐	☐	☐	☐	☐
- Patient(s) or consumer group	☐	☐	☐	☐	☐	☐	☐

20. How often do the following issues impact your organization/group's decision-making process of whether to fund or conduct an update? Please tick one answer for each reason

	Always	Often	Sometimes	Seldom	Never	Not sure	Not applicable
- Cost utility of updating	☐	☐	☐	☐	☐	☐	☐
- Impact on policy	☐	☐	☐	☐	☐	☐	☐
- Impact on clinical practice	☐	☐	☐	☐	☐	☐	☐
- Political context	☐	☐	☐	☐	☐	☐	☐
- Current public controversy/interest	☐	☐	☐	☐	☐	☐	☐
- Burden of illness/costs of disease	☐	☐	☐	☐	☐	☐	☐
- Organization/group credibility of being current	☐	☐	☐	☐	☐	☐	☐
- Anticipated change in effect size or precision	☐	☐	☐	☐	☐	☐	☐
- To address a quality gap in care	☐	☐	☐	☐	☐	☐	☐

21. Please list any other additional issues that may factor into your organization/group determining WHEN to update a systematic review using the box provided below.
If there are no additional issues that you are aware of, please leave blank.

22. How would you best describe the methods or strategies used by your organization/group to determine HOW to conduct an update of a systematic review? Please tick one of the following options

- ☐ Formal method(s) used
- ☐ Informal method(s) used
- ☐ No method(s) used
- ☐ Not sure

23. Please describe the formal/informal method(s) in the box provided below (if applicable):

24. On average, which of the following best estimates the time your organization/group typically expends on implementing an updating strategy?
Please tick one of the following options

- ☐ 0-3 months
- ☐ 4-6 months
- ☐ 7-9 months
- ☐ 10-12 months
- ☐ 13-15 months
- ☐ 16-18 months
- ☐ 19-21 months
- ☐ 21-24 months

- >24 months
- Not sure
- Not applicable

25. On average, what is your organization/group's estimated budget for implementing an updating strategy? (U.S. Dollars) Please tick one of the following options

- <$10,000
- $10,000 - $20,000
- $21,000 - $30,000
- $31,000 - $40,000
- $41,000 - $50,000
- $51,000 - $60,000
- $61,000 - $70,000
- $71,000 – $80,000
- $81,000 – $90,000
- $91,000 - $100,000
- $100,001 - $200,000
- >$200,000
- Not sure
- Not applicable

You have reached the half-way mark of this survey.

26. How often has your organization/group been involved in the following scenarios? Please tick one answer for each activity.

	Always	Often	Sometimes	Seldom	Never	Not sure	Not applicable
- A partial update involving only certain sections of a review	☐	☐	☐	☐	☐	☐	☐
- A full update of all sections	☐	☐	☐	☐	☐	☐	☐
- An entirely new review upon updating	☐	☐	☐	☐	☐	☐	☐
- Knowing a systematic review is out of date but unable to commence updating due to lack of resources (e.g. funding, personnel, time etc.)	☐	☐	☐	☐	☐	☐	☐

27. When you do an update, how often is your organization/group able to draw on the same people who did the original review?
Please tick one of the following options

- Always
- Often
- Sometimes
- Seldom
- Never
- Not sure

☐ Not applicable

28. How frequently does your organization/group use the following procedures when updating? Please tick one answer for each reason

	Always	Often	Sometimes	Seldom	Never	Not sure	Not applicable
- Time-specific approach (pre-set updating frequency)	☐	☐	☐	☐	☐	☐	☐
- Editorial strategy with an algorithm of administrative actions	☐	☐	☐	☐	☐	☐	☐
- Bibliometric database entry-date searching	☐	☐	☐	☐	☐	☐	☐
- Cumulative meta-analysis (CMA) or its' extensions (e.g., recursive; with sequential boundaries; or using cumulative slope)	☐	☐	☐	☐	☐	☐	☐
- Barrowman's diagnostic test (i.e., identifying 'null' meta-analyses ready for updating)	☐	☐	☐	☐	☐	☐	☐

29. Please list any other additional procedures used during the updating process or list any relevant working documents your organization has developed detailing these procedures in the box provided below.

If there are no new additional strategies used during the update process that you are aware of, please leave blank.

30. How often has your organization/group been involved in updating systematic reviews done by others? Please tick one of the following options

☐ Always ☐ Often ☐ Sometimes ☐ Seldom ☐ Never ☐ Not sure ☐ Not applicable

31. When updating reviews done by others, how often is your organization/group involved in the following: Please tick one answer for each activity

	Always	Often	Sometimes	Seldom	Never	Not sure	Not applicable
- Revising the original search strategy	☐	☐	☐	☐	☐	☐	☐
- Revising the eligibility criteria	☐	☐	☐	☐	☐	☐	☐
- Re-abstracting data from previously reviewed studies	☐	☐	☐	☐	☐	☐	☐
- Re-assessing quality of previously reviewed studies	☐	☐	☐	☐	☐	☐	☐

32. To what extent are the following characteristics of original reviews perceived as barriers to updating reviews done by others? Please tick one answer for each reason

	Serious Barrier	Moderate Barrier	Minor Barrier	Not a Barrier	Not sure	Not applicable
Perceived need to change the original search strategy	☐	☐	☐	☐	☐	☐
Perceived need to change the original screening criteria	☐	☐	☐	☐	☐	☐
Perceived need to redo the data abstraction	☐	☐	☐	☐	☐	☐
Perceived need to re-assess study quality	☐	☐	☐	☐	☐	☐

33. How often does your organization/group identify recent relevant literature published after the date of the last search but before completion of the final draft?
Please tick one of the following options

- ☐ Always
- ☐ Often
- ☐ Sometimes
- ☐ Seldom
- ☐ Never
- ☐ Not sure
- ☐ Not applicable

34. **When recent relevant literature is identified as being published after the date of the last search but before completion of the final draft, how does your organization/group usually incorporate this new evidence?**

Please check all that apply

- ☐ As an addendum in the review
- ☐ As a formal revision to the analysis
- ☐ Not sure
- ☐ Not applicable
- ☐ Other

35. To what extent would you consider the following as general barriers to your organization/group updating systematic reviews?
Please tick one answer for each reason

	Serious Barrier	Moderate Barrier	Minor Barrier	Not a Barrier	Not sure	Not applicable
- Limited updating methodologies	☐	☐	☐	☐	☐	☐
- Limited funding/resources	☐	☐	☐	☐	☐	☐
- Work redundancy of updating (reviewer motivation)	☐	☐	☐	☐	☐	☐
- Limited academic credit for updating	☐	☐	☐	☐	☐	☐

work						
- Limited journal publishing formats	☐	☐	☐	☐	☐	☐
- Having to update reviews done by others	☐	☐	☐	☐	☐	☐

36. To what extent does your organization/group agree or disagree with harmonizing updating efforts across organizations or groups that fund or conduct systematic reviews? Please tick one of the following options

- ☐ Strongly agree
- ☐ Somewhat agree
- ☐ Neither agree nor disagree
- ☐ Somewhat disagree
- ☐ Strongly disagree
- ☐ No opinion

37. To what extent would you consider the following as benefits to your organization/group participating in international harmonization efforts for updating systematic reviews? Please tick one answer for each reason

	Major benefit	Moderate benefit	Minor benefit	No benefit	Not sure	Not applicable
- Access to new information, ideas, materials or other resources	☐	☐	☐	☐	☐	☐
- Potential to minimize duplication of services	☐	☐	☐	☐	☐	☐
- Use of existing resources more efficiently	☐	☐	☐	☐	☐	☐
- Ability to address issues beyond a single organization/group's domain	☐	☐	☐	☐	☐	☐
- Shared responsibility across organization/group's for complex or controversial issues	☐	☐	☐	☐	☐	☐

38. To what extent would you consider the following as barriers to your organization/group participating in international harmonization efforts for updating systematic reviews? Please tick one answer for each reason

	Serious Barrier	Moderate	Minor	Not a Barrier	Not sure	Not applicable
- Diversion of an organization/group's human resources	☐	☐	☐	☐	☐	☐
- Diversion of an organization/group's	☐	☐	☐	☐	☐	☐

funding resources						
- Diversion from an organization/group's research mandate	☐	☐	☐	☐	☐	☐
- Perceived delays in working across organizations/groups	☐	☐	☐	☐	☐	☐

39. To what extent would your organization/group agree or disagree with development of a central registry of systematic reviews including existing systematic reviews and protocols (i.e., similar to efforts within the clinical trials community)? Please tick one of the following options

- Strongly agree
- Somewhat agree
- Neither agree nor disagree
- Somewhat disagree
- Strongly disagree
- No opinion

40. Does your organization/group encourage the following activities when funding or conducting systematic reviews?
Please tick one answer for each activity

	Yes	No	Not sure
- Inclusion of a prediction for updating in the text of original reviews	☐	☐	☐
- Withdrawal of systematic reviews from circulation when assessed as out of date	☐	☐	☐
- Formal retirement of systematic reviews when deemed no longer in need of further investigation	☐	☐	☐

41. Which of the following best describes your organization/group?
Please tick one of the following options

- Academic Institution
- National Government Agency
- Regional/Local Government Agency
- Private Organization (Industry)
- Medical Specialty Society
- Disease Specific Society
- Managed Care Organization
- Other (please specify)

42. What is your organization/group's primary funding structure?
Please tick one of the following options

- For profit
- Not for profit

43. What type(s) of funding does your organization receive?
Please tick all that apply

- Industry/Private Sector
- Government (Infrastructure; grants)

- Non-profit (Academic; non-governmental organizations)
- Endowment fund
- Internal
- Other (please specify)

44. What is your organization/group's total estimated proportion of work expended annually on the following research synthesis areas?
Please tick one answer for each activity

	Original systematic reviews	Updating of systematic reviews	Clinical practice guidelines (CPGs)	Health technology assessments (HTAs)
<10%	☐	☐	☐	☐
11-20%	☐	☐	☐	☐
21-30%	☐	☐	☐	☐
31-40%	☐	☐	☐	☐
41-50%	☐	☐	☐	☐
51-60%	☐	☐	☐	☐
61-70%	☐	☐	☐	☐
71-80%	☐	☐	☐	☐
81-90%	☐	☐	☐	☐
91-100%	☐	☐	☐	☐
Not sure	☐	☐	☐	☐
Not applicable	☐	☐	☐	☐

45. What is your organization/group's estimated budget generated/expended annually on the following research synthesis areas in US Dollars?
Please tick one option for each activity

	Original systematic reviews	Updating of systematic reviews	Clinical practice guidelines (CPGs)	Health technology assessments (HTAs)
<10%	☐	☐	☐	☐
$11,000-40,000	☐	☐	☐	☐
$41,000-70,000	☐	☐	☐	☐
$71,000-$100,000	☐	☐	☐	☐
$100,000-200,000	☐	☐	☐	☐
$200,000-500,000	☐	☐	☐	☐
$500,000-1,000,000	☐	☐	☐	☐
>$1,000,000	☐	☐	☐	☐
Not sure	☐	☐	☐	☐
Not applicable	☐	☐	☐	☐

46. In what capacity have you previously participated in a systematic review?
Please tick all that apply

- Lead author of a systematic review
- Co-author of a systematic review
- Information specialist
- Statistician
- Methodologist/Epidemiologist
- Clinical expert
- Editor
- Project Manager/Coordinator
- None
- Other research capacity (please specify)

47. Please describe your training in systematic reviews.
Please tick all that apply

- University-level training or research in systematic reviews
- Continuing education course(s) in systematic reviews
- Workshop(s) in systematic reviews
- Lecture(s) in systematic reviews
- No training in systematic reviews
- Other (Please specify)

48. How often do you utilize evidence from systematic reviews in the following capacities?
Please tick one of the following options

	Daily	Weekly	Monthly	Never	Not applicable
To inform clinical practice	☐	☐	☐	☐	☐
To inform policy-making	☐	☐	☐	☐	☐
To inform funding decisions	☐	☐	☐	☐	☐
For research purposes	☐	☐	☐	☐	☐

49. Would you like to receive a summary of this research when it is available?
Please tick one of the following options

- Yes
- No

50. May we contact your organization/group within the next 12 months for more in depth information on its updating experiences? Please tick one of the following options

- Yes
- No

---

51. Please enter your contact information below. This will remain confidential and will only be viewed by the primary researcher.

Responses are required for each field below:
First & Last Name:
Email address:
Name of Organization/Group:
Title or position:
Degrees held:
Numbers of Yrs with the Organization/Group:
Street Address of the Organization/Group:
City:
Province/State/Region:
Country:
Postal/Zip Code:
Website address of your organization:

You have now completed the survey! As a token of our appreciation for your time, an Amazon voucher will be sent to you shortly.

On behalf of the University of Ottawa EPC, we thank you!