Methods Research Report

Selective Outcome Reporting as a Source of Bias in Reviews of Comparative Effectiveness

Agency for Healthcare Research and Quality
Advancing Excellence in Health Care • **www.ahrq.gov**

Methods Research Report

Selective Outcome Reporting as a Source of Bias in Reviews of Comparative Effectiveness

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. 290-2007-10057-I

Prepared by:
Oregon Evidence-based Practice Center
Portland, OR

Investigators:
Susan L. Norris, M.D., M.P.H, M.Sc.
Haley K. Holmer, M.P.H.
Lauren A. Ogden, B.A.
Rongwei Fu, Ph.D.
Ahmed M. Abou-Setta, M.D., Ph.D.
Meera S. Viswanathan, Ph.D.
Melissa L. McPheeters, Ph.D., M.P.H.

AHRQ Publication No. 12-EHC110-EF
August 2012

This report is based on research conducted by the Oregon Evidence-based Practice Center (EPC) under contract to the Agency for Healthcare Research and Quality (AHRQ), Rockville, MD (Contract No. 290-2007-10057-I). The findings and conclusions in this document are those of the authors, who are responsible for its contents; the findings and conclusions do not necessarily represent the views of AHRQ. No statement in this report should be construed as an official position of AHRQ or of the U.S. Department of Health and Human Services.

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

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

Persons using assistive technology may not be able to fully access information in this report. For assistance, contact EffectiveHealthCare@ahrq.gov.

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

Suggested citation: Norris SL, Holmer HK, Ogden LA, Fu R, Abou-Setta AM, Viswanathan MS, McPheeters ML. Selective Outcome Reporting as a Source of Bias in Reviews of Comparative Effectiveness. (Prepared by the Oregon Evidence-based Practice Center under Contract No. 290-2007-10057-I.) AHRQ Publication No. 12-EHC110-EF. Rockville, MD: Agency for Healthcare Research and Quality. August 2012. www.effectivehealthcare.ahrq.gov/reports/final.cfm.

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 and strategies. 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 improve the scientific rigor of these evidence reports, AHRQ supports empiric research by the EPCs to help understand or improve complex methodologic issues in systematic reviews. These methods research projects are intended to contribute to the research base in and be used to improve the science of systematic reviews. They are not intended to be guidance to the EPC program, although may be considered by EPCs along with other scientific research when determining EPC program methods guidance.

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. The reports undergo peer review prior to their release as a final report.

We welcome comments on this Methods Research Project. 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 email to epc@ahrq.hhs.gov.

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

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

Stephanie Chang, M.D., M.P.H.
Director
Evidence-based Practice Program
Center for Outcomes and Evidence
Agency for Healthcare Research and Quality

Mary Nix, M.S., M.T. (A.S.C.P.) S.B.B.
Task Order Officer
Center for Outcomes and Evidence
Agency for Healthcare Research and Quality

Acknowledgments

The authors gratefully acknowledge Thomas A. Trikalinos, M.D., Ph.D., Sasha Walia, M.P.A., Katie Reitel, B.A., and Leah Williams, B.S., for their technical contributions to this project.

Peer Reviewers

Lisa Bero, Ph.D.
Professor
University of California, San Francisco
San Francisco, CA

An-Wen Chan M.D., D.Phil.
Assistant Professor
University of Toronto
Toronto, Canada

Davina Ghersi, M.P.H., Ph.D.
World Health Organization
Geneva, Switzerland

Yoon Loke, M.D., M.B.B.S.
Clinical Senior Lecturer
University of East Anglia
Norwich, U.K.

Erik von Elm, M.D., M.Sc.
University Hospitals of Lausanne
Institut of Social and Preventive Medicine
Lausanne, Switzerland

Paula Williamson, Ph.D.
Professor
University of Liverpool
Liverpool, U.K.

Deborah Zarin, M.D.
Director of ClinicialTrial.gov
National Library of Medicine
National Institutes of Health

Selective Outcome Reporting as a Source of Bias in Reviews of Comparative Effectiveness

Structured Abstract

Objectives. The objectives of this exploratory study were to: **(**1) describe the frequency of selective outcome reporting (SOR) and selective analysis reporting (SAR) in randomized controlled trials (RCTs) included in reviews of comparative effectiveness for outcomes of benefit; (2) explore potential predictors for SOR and SAR; and (3) assess the reliability and validity of the Outcome Reporting Bias in Trials (ORBIT) classification system for missing or incomplete outcome reporting.

Data Sources and Methods. We selected three comparative effectiveness reviews (CERs) funded by the Agency for Healthcare Research and Quality that included drug–drug comparisons. Within each CER, we then specified one outcome that fulfilled explicit criteria (the "index outcome") and examined the RCTs in the CER that reported that outcome. We then searched trial registries for study registration information and results for each RCT. Using available registry information to complement information in the methods section of the publication, we determined the frequency of SOR and SAR, and we examined prespecified predictors of SOR and SAR. Lastly, using the ORBIT classification of SOR, we attempted to examine the inter-rater reliability of ORBIT and its validity, comparing information contained within the publication to assessments of SOR, using the additional information obtained from trial registries.

Results. RCTs published in 2005 or later and reporting the index outcome were not consistently listed in trial registries, with 29 percent, 67 percent, and 75 percent of trials registered for each of the three CERs. In addition, publications did not consistently report trial registration. Results were infrequently listed in ClinicalTrials.gov, even after 2008, when reporting became mandatory for certain types of trials. Trial registration frequently occurred after the study was completed (in 25 percent, 50 percent, and 42 percent of trials in each of the three CERs). Changes occurred in the specification of the index outcome in the registry in 42 percent and 17 percent of trials in two CERs (the index outcome in the third CER was never mentioned in the registry). We did not find the ORBIT classification tool particularly useful: it was difficult to implement, and the nine classes were difficult to reliably distinguish. In addition, ORBIT classes did not describe a type of SOR and SAR that we frequently encountered: the addition of outcomes measures, subgroups, and other analyses to published results that were not prespecified in the publication's methods section or listed in the registry. Finally, trial registries were of little use in identifying SOR unless trial results were listed in the registry and of no use in identifying SAR.

Conclusions. We identified numerous challenges in identifying and characterizing SOR and SAR in this pilot study of three CERs. Existing tools were suboptimal: ORBIT does not encompass the type of SOR and SAR where results in the publication were not prespecified in the methods section or in the registry. The design of our study (focusing on RCTs with results in the CER) precluded identifying certain types of SOR where the outcomes were not reported at all

in the study. The presentation and content of ClinicalTrials.gov could be improved to better assist the systematic reviewer in identifying potential SOR and SAR. Further research is needed to develop efficient, tailored approaches to identifying and characterizing SOR and SAR in trials.

Contents

Tables

Figures

Appendixes

Background

Reporting biases arise when the dissemination of research findings is influenced by the nature and direction of the results,[1] and can arise from processes acting within a study or at the level of the whole study. Within studies, researchers may report their findings selectively—choosing to report selected outcomes and analyses based on the results. Reporting bias can thus result from selective outcome reporting (SOR), wherein only a subset of the original outcomes measured and analyzed in a study are fully reported based on the magnitude of the treatment effect or the statistical significance of selected outcomes.[2]

Kirkham and colleagues describe three main types of SOR:[3] selective reporting of an entire study outcome (i.e., analyzed outcomes are not reported); selective reporting of a specific outcome (e.g., selected followup intervals), and incomplete reporting of a specific outcome (e.g., incomplete reporting of nonsignificant p values, such as $p>0.05$). SOR can result in outcome reporting bias (ORB), which is the bias produced from choosing which outcomes to publish based on the results.[2]

Reporting bias arising from within-study processes can also result from the selection of analyses for reporting (SAR), which can lead to analysis reporting bias (ARB). Examples of SAR include selective reporting of data on subgroups, presentation of adjusted rather than unadjusted analyses, selection of as-treated rather than intention-to-treat analyses, selective approaches to the handling of missing data, choosing to analyze continuously measured variables categorically (outcomes or predictors in adjusted models), and choice of cut-point values to define categorical variables.[4]

The high prevalence of SOR and SAR among primary studies is well documented. This research has been done almost exclusively in randomized controlled trials (RCTs), usually by comparing study protocols submitted to regulatory or funding agencies with published outcomes.[5-7] In a systematic review of five such cohorts (four of which contained only RCTs), Dwan and colleagues[6] reported that changes in prespecified outcomes occurred that were not documented in protocol amendments in 40 to 62 percent of studies, where there was at least one primary outcome that was changed, introduced, or omitted between the protocol and the publication. In addition, statistically significant outcomes had a higher odds of being fully reported compared to nonsignificant outcomes (range of odds ratios [OR], 2.2 to 4.7), suggesting ORB as well as SOR.

There are few studies on the prevalence of SOR and SAR among trials included in systematic reviews, and little is known about the effects of selective reporting on effect estimates and conclusions in such reviews. Kirkham and colleagues[3] compared effect estimates reported in meta-analyses to estimates obtained with sensitivity analyses estimating the same effects without SOR (using the maximum bias bound approach[8]) for a sample of new systematic reviews published in the Cochrane Library. Of 81 reviews with a single meta-analysis of the review primary outcome, 52 (64 percent) included one or more RCTs with a high suspicion of ORB. Of 25 reviews that could be assessed, the median percentage change in treatment effect between the reported effect and the estimated effect without SOR was 39 percent (interquartile range 18 to 67 percent). Of 42 meta-analyses with statistically significant results, 19 percent became non-significant after adjustment for ORB and 26 percent overestimated the treatment effect by 20 percent or more. Hart and coauthors[9] reanalyzed meta-analyses of drug efficacy and harms, adding unpublished data from the U.S. Food and Drug Administration (FDA), and reported a change in the assessment of efficacy of the drug in 92 percent of the meta-analyses.

Kirkham and colleagues[3] developed a classification system for SOR, called Outcome Reporting Bias in Trials (ORBIT) (Table 1). This nine-category assessment tool is based on information in the trial publication(s) only and not on other information such as that contained in trial registries. This system focuses on outcomes that are missing or incompletely reported in reports of RCTs, and differentiates types of SOR based on the assessor's certainty about whether the outcome was measured and analyzed, and the potential reasons for missing data.

In the design stage of a study, outcomes can be selected based on anticipated results, and these selected outcomes can be then specified in the study protocol. By definition this is not SOR as the selection of outcomes is not based on actual results; however this approach to design of primary studies can ultimately lead to biased results and conclusions in systematic reviews.

Publication bias, whereby an entire study is not published because of the nature or direction of the results,[10, 11] is also an important issue for systematic reviewers. Statistically significant results are more likely to be published than studies with "negative" or "null" findings,[12] and positive findings are more likely to be published rapidly,[13, 14] in English, with multiple companion papers, in high impact journals, and to be cited by others.[5, 15] In this report we focus exclusively on the less well studied and recognized issues of within-study selective reporting, specifically SOR and SAR, and do not examine publication bias.

Systematic reviewers should assess the risk of all potential biases in included primary studies. Given that there are emerging data suggesting the presence of SOR and SAR, systematic reviewers need to consider the potential bias due to missing outcomes or analyses among the primary studies included in a review. In addition, review authors need to consider how SOR and SAR might affect the direction, magnitude, and precision of pooled effect estimates, as well as the conclusions about both benefits and harms in systematic reviews.

There are no data that we are aware of on the effects of SOR and SAR in reviews of comparative effectiveness and it is possible that selective reporting (SOR and/or SAR) has different frequencies and implications across various types of systematic reviews, interventions, and outcomes. For example, the availability of protocols may vary among types of interventions (e.g., drug vs. behavioral therapy) and studies (e.g., effectiveness vs. efficacy). In addition, some of the characteristics of comparative effectiveness research may affect the frequency and impact of SOR and SAR: comparative effectiveness reviews are more likely to include subjective measures of patient-important outcomes (e.g., symptoms, quality of life), head-to-head rather than placebo-controlled studies, and heterogeneity of populations and interventions. Evidence on selective reporting across various study designs and outcomes may assist in the interpretation of summary effect measures and conclusions in reviews of comparative effectiveness.

The registration of studies, particularly RCTs, is an important tool for identifying all studies related to key question in a comparative effectiveness review (CER). Registries are also a potential tool for assessing SOR and SAR. In the United States, the U.S. Food and Drug Administration Modernization Act of 1997 called for the creation of ClinicalTrials.gov and mandated registration of all efficacy drug trials for serious or life-threatening diseases and conditions conducted under FDA Investigational New Drug Application regulations.[16] Each record in ClinicalTrials.gov includes summary information on the study protocol, patient recruitment status, and the location of the study site. Beginning in September, 2008, the FDA requires that results also be reported in clinicalTrials.gov, although some exceptions are permitted.[17]

The World Health Organization (WHO) initiated a policy in 2006 requiring trial registration of all medical studies that test treatments on patients or healthy volunteers.[18] WHO developed

the International Clinical Trials Registry Platform (ICTRP), a global initiative that aims to make information about all clinical trials involving humans publicly available (www.who.int/ictrp/network/primary/en/index.html).[18] The ICTRP operates a Search Portal, which provides access to information about ongoing and completed clinical trials from a number of different trial registries (See Appendix A).

Table 1. The Outcome Reporting Bias In Trials (ORBIT) study classification system for missing or incomplete outcome reporting in reports of randomized trials[3]

Category	Description	Level of Reporting	Risk of Bias
	Clear that the outcome was measured and analyzed		
A	Trial report states that outcome was analyzed but only reports that result was not significant (typically stating p>0.05)	Partial	High risk
B	Trial report states that outcome was analyzed but only reports that result was significant (typically stating p<0.05)	Partial	No risk
C	Trial report states that outcome was analyzed but insufficient data were presented for the trial to be included in meta-analysis or to be considered to be fully tabulated	Partial	Low risk
D	Trial report states that outcome was analyzed but no results reported	None	High risk
	Clear that the outcome was measured		
E	Clear that outcome was measured but not necessarily analyzed. Judgment says likely to have been analyzed but not reported because of non-significant results	None	High risk
F	Clear that outcome was measured but not necessarily analyzed. Judgment says unlikely to have been analyzed but not reported because of non-significant results	None	Low risk
	Unclear whether the outcome was measured		
G	Not mentioned but clinical judgment says likely to have been measured and analyzed but not reported on the basis of non-significant results	None	High risk
H	Not mentioned but clinical judgment says unlikely to have been measured at all	None	Low risk
	Clear that the outcome was not measured		
I	Clear that outcome was not measured	NA	No risk

Abbreviations: NA, not applicable.

This table presents the categories of ORBIT that Kirkham and colleagues proposed,[3] including their categorization of level of reporting their assessment of the risk for bias. Reproduced with permission.

Objectives

This exploratory study set about to examine the frequency and effect of reporting biases, specifically SOR and SAR, in reviews of comparative effectiveness. This work focused specifically on using trial registries as a potential tool for assessing SOR and SAR. The goal of this study was to inform ongoing work in AHRQ's Evidence-based Practice Center (EPC) program to develop valid and efficient approaches and procedures for identifying SOR and SAR in studies included in systematic reviews, and to assess the risk of bias due to missing data in CERs.

We defined outcomes rather broadly, in order to encompass a change in outcome specification (e.g., followup interval or continuous to categorical variable) in our examination of SOR. We also wanted to examine the prevalence of the addition of outcomes (that were not prespecified) to a publication. We did not focus on the type of SOR where outcomes were missing completely from a publication which likely could or should have reported them, as exploration of that type of SOR would have markedly increased the scope of our project.

The specific objectives of this task order were to:

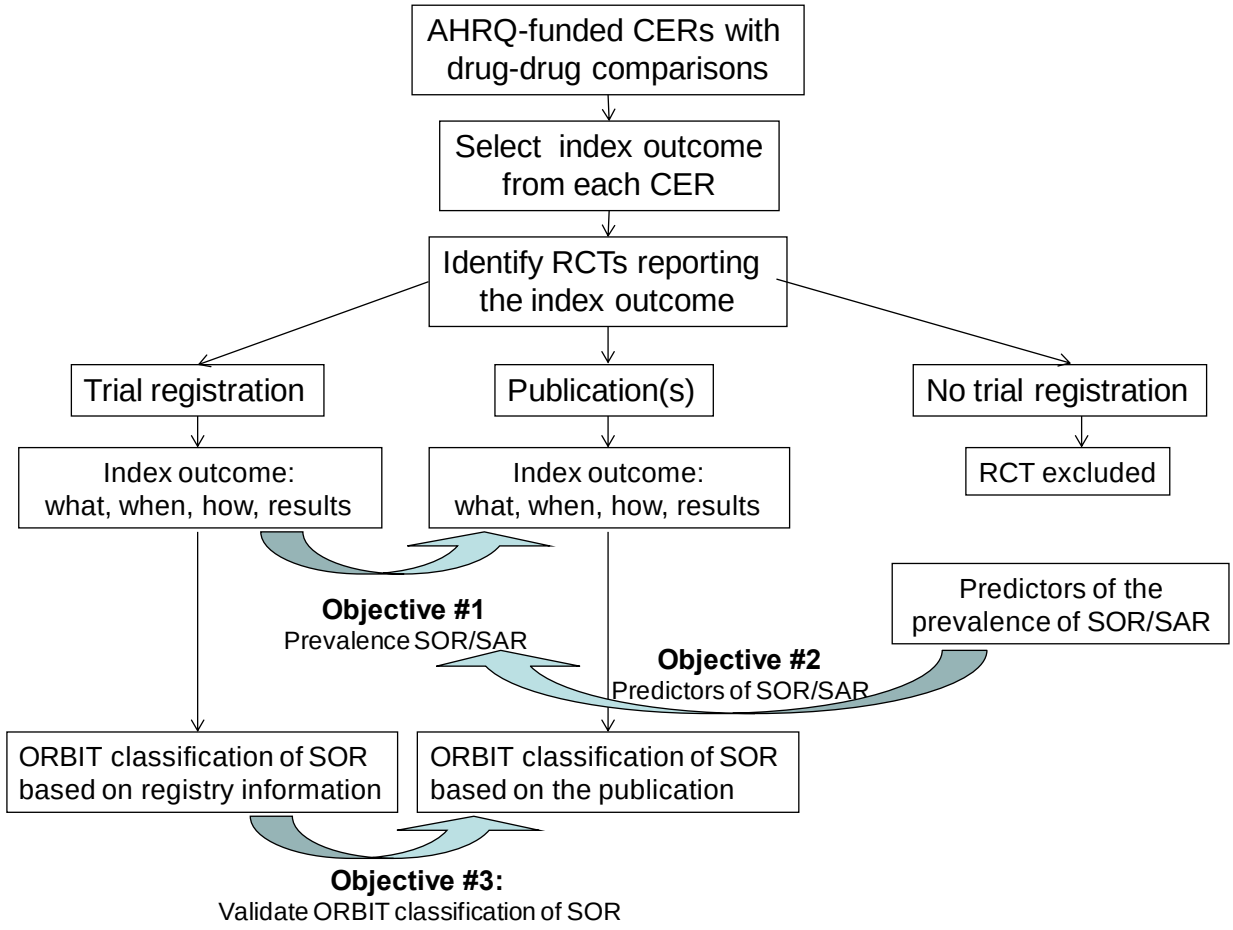
1. describe the frequency of SOR and SAR within primary studies included in reviews of comparative effectiveness for outcomes of benefit;
2. explore potential predictors of SOR and SAR in RCTs; and
3. assess the reliability and validity of the ORBIT[3] classification system for missing or incomplete outcome reporting and the ORBIT assessment of the risk of bias associated with different types of SOR.

Methods

Overview of Methods

Figure 1 provides an overview of the methods used in this project. Among CERs funded by AHRQ, we identified those that fulfilled our inclusion criteria. Within each of these CERs, we specified one outcome that fulfilled explicit criteria (the index outcome). For this pilot study we selected three CERs to examine in detail. RCTs examining the index outcome were then identified, and trial registries were searched for study registration and protocols. Focusing on RCTs that were listed in a trial registry, we then used registry information, along with the methods section of the publication, to determine the frequency of various types of SOR and SAR (Objective #1). For Objective #2, we examined prespecified predictors of the existence of SOR and SAR in RCTs. Lastly, using the ORBIT classification of SOR,[3] we attempted to examine the inter-rater reliability of ORBIT and its validity, comparing SOR assessments using information contained within the publication to assessments of SOR using the additional information in trial registries (Objective #3).

Figure 1. Overview of report methodology



Abbreviations: AHRQ, Agency for Healthcare Quality and Research; CER, comparative effectiveness review; RCT, randomized controlled trial; SAR, selective analysis reporting; SOR, selective outcome reporting.

Selection of the Cohort of Comparative Effectiveness Reviews

To explore SOR and SAR, we selected a cohort of CERs that fulfilled the criteria outlined in Table 2. We desired a cohort that was relevant to the AHRQ Effective Health Care Program, was feasible to examine with the resources allocated to this project, would provide results applicable to future AHRQ CERs, would contribute to the existing methodological literature on SOR and SAR, and would facilitate the development of specific guidance on the detection and implications of SOR and SAR for systematic reviewers in the Evidence-based Practice Center (EPC) Program. Of the 26 currently published CERs, 15 fulfilled these inclusion criteria (Table 3).

Table 2. Inclusion criteria for the AHRQ-funded systematic reviews examined in this report

General Criteria	Specific Criteria	Rationale	Number of Reports Fulfilling the Inclusion Criteria[a]
Type of review	Comparative effectiveness review	Reviews of comparative effectiveness are used for clinical and public health decision-making, thus understanding SOR and SAR for this type of systematic review is important. Technical briefs were excluded as they examine devices or diagnostic tests and include a very small number of RCTs.	26
Review status	Final	Final reports were examined in order to avoid reports that may be revised in the near future, and to include reports that have been reviewed and incorporated input from peer reviewers, public comments, and the AHRQ Associate Editor.	25
Focus of the CER	The focus must be pharmacotherapy with drug-drug or drug-placebo comparators.	Methodology of drug reviews is more advanced in general than reviews of devices, procedures, prognosis, diagnosis, or epidemiology. RCTs are likely to be included, and some sources of clinical and methodological heterogeneity may be minimized by examining only drug interventions.	15
Design of included studies	The CER must contain 1 or more RCTs.	Only RCTs were examined as we sought to identify study registration and studies with other designs are much less likely to be registered.	15
Outcomes	The CER must examine benefits.	Benefit outcomes may be more likely to be delineated in a trial registry than are specific harms. SOR of harms is also an important issue, but resource constraints did not permit us to explore harms in this project.	15

Abbreviations: AHRQ, Agency for Healthcare Research and quality; CERs, comparative effectiveness reviews; NA, not applicable; RCT, randomized controlled trial; SAR, selective outcome reporting; SOR, selective outcome reporting.

[a] Reports were identified on April 12, 2011; the number of AHRQ CERs changes as additional reports are published. On that date, there were 26 CERs available on the AHRQ Website.

Table 3. Index outcome for each comparative effectiveness review

Comparative Effectiveness Review	Index Outcome	Number of RCTs[b]
1. Oral Diabetes Medications for Adults with Type 2 Diabetes. An Update[19] [a]	A1c	24
2. Effectiveness of Recombinant Human Growth Hormone (rhGH) in the Treatment of Patients With Cystic Fibrosis[20]	Pulmonary function testing, including both forced expiratory volume in one second (FEV1) and forced vital capacity (FVC)	1
3. Comparative Effectiveness of In-Hospital Use of Recombinant Factor VIIa for Off-Label Indications vs. Usual Care[21]	Mortality	4
4. Comparative Effectiveness of Angiotensin Converting Enzyme Inhibitors or Angiotensin II Receptor Blockers Added to Standard Medical Therapy for Treating Stable Ischemic Heart Disease[22]	Mortality	1
5. Comparative Effectiveness of Medications To Reduce Risk of Primary Breast Cancer in Women[23]	Invasive and non-invasive breast cancer	5
6. Comparative Effectiveness of Lipid-Modifying Agents[24] [a]	All-cause mortality	14
7. Comparative Effectiveness, Safety, and Indications of Insulin Analogues in Premixed Formulations for Adults With Type 2 Diabetes[25]	A1c	11
8. Comparative Effectiveness of Treatments To Prevent Fractures in Men and Women With Low Bone Density or Osteoporosis[26] [a]	Fracture reduction, including both vertebral and extremity/hip/etc	12
9. Comparative Effectiveness of Drug Therapy for Rheumatoid Arthritis and Psoriatic Arthritis in Adults[27]	American College of Rheumatology 50	1
10. Comparative Effectiveness of Angiotensin-Converting Enzyme Inhibitors (ACEIs) and Angiotensin II Receptor Antagonists (ARBs) for Treating Essential Hypertension[28]	Blood pressure control	2
11. Comparative Effectiveness of Second-Generation Antidepressants in the Pharmacologic Treatment of Adult Depression[29]	Response in treating depressive symptoms (according to various depression rating scales)	0
12. Efficacy and Comparative Effectiveness of Off-Label Use of Atypical Antipsychotics[30]	Obsessive-compulsive disorder: Yale-Brown Obsessive-Compulsive Scale	1
13. Comparative Effectiveness and Safety of Analgesics for Osteoarthritis[31]	Pain relief	0
14. Comparative Effectiveness of Epoetin and Darbepoetin for Managing Anemia in Patients Undergoing Cancer Treatment[32]	Hemoglobin levels	2
15. Comparative Effectiveness of Management Strategies for Gastroesophageal Reflux Disease[33]	Complete symptom relief at 4 and 8 weeks; time to complete resolution of symptoms	1

Abbreviations: A1c, hemoglobin A1c; RCTs, randomized controlled trials.

[a] Comparative Effectiveness Reviews selected for further evaluation of SOR and SAR .

[b] Number of RCTs that reported the index outcome, were published in or after 2005, and which were registered.

Identification of an Index Outcome for Each Included Comparative Effectiveness Review

We next identified a single effectiveness or efficacy outcome for each of the 15 CERs according to the criteria in Figure 2. Where available, the primary outcome of the CER was used as the index outcome for that CER. More commonly, the CER did not specify a primary outcome and two authors of this methods report independently selected one outcome as the index outcome based on the prespecified criteria. When differences between the two authors occurred, consensus was achieved through discussion.

The index outcome selected for each CER must be one that is consistently measured using the same technique and measurement scale because it would be impossible to compare outcomes measured in different ways across studies and to attribute missing information to SOR or SAR rather than to differences in measurement approach. We also selected outcomes for which a meta-analysis was presented in the CER as initially we planned to compare pooled estimates in the CER with estimates using imputation for missing outcomes (Objective #3).

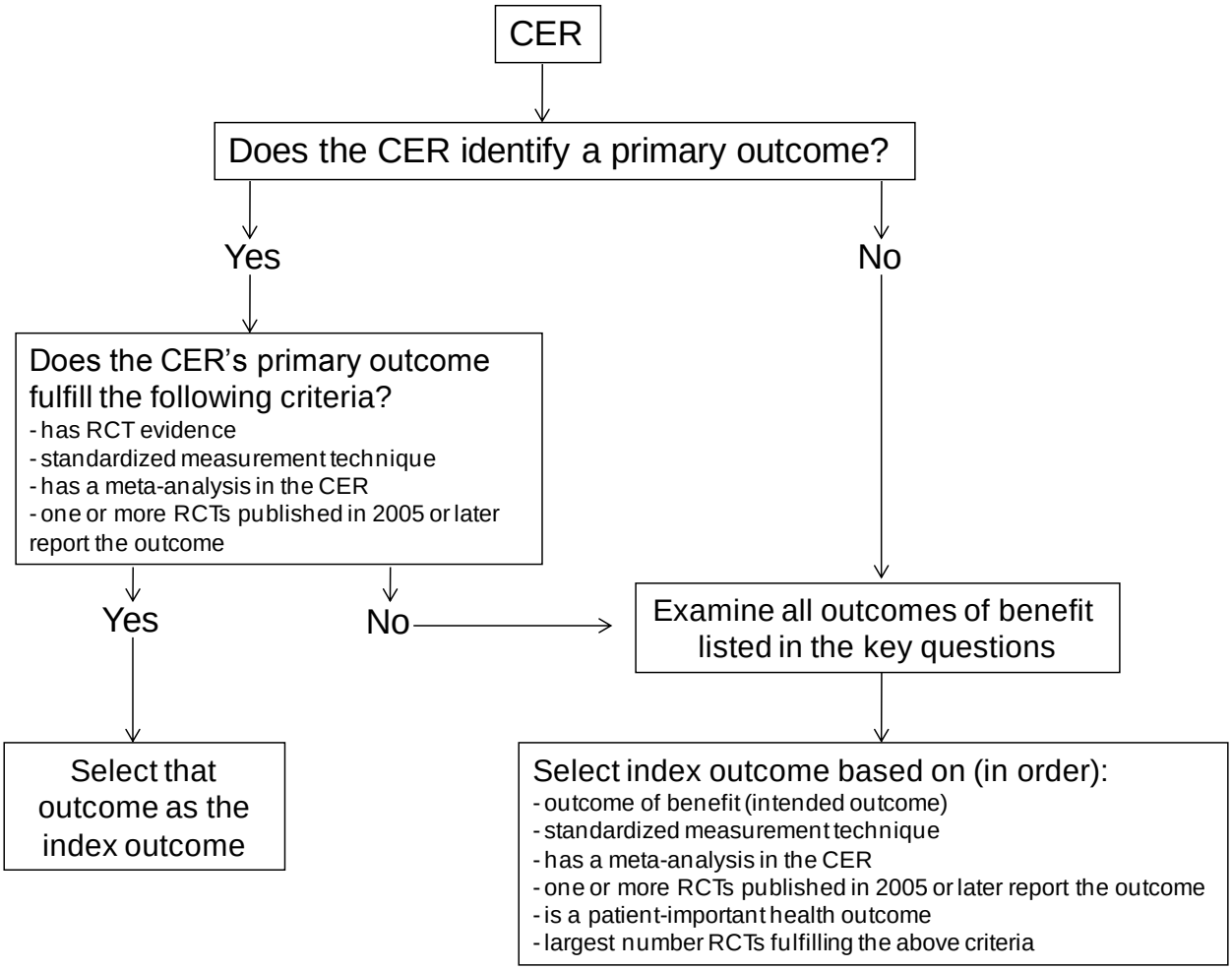
We specified our index outcome at the level of a specific measurement (e.g., hemoglobin A1c [A1c]). We did not specify the metric (e.g., change from baseline) or the method of aggregation (e.g., continuous or categorical such as proportion of persons with A1c <7.0 percent). Thus our index outcomes were specified at level 2 (specific measurement) in the categorization proposed by Zarin and colleagues.[34] We defined our index outcome with this level of specificity in order to explore the prevalence of changes in specification of outcomes within publications and between registry information and trial publications. For example, for the CER on oral hypoglycemic agents, we specified "A1c" as our index outcome and then looked for changes in followup intervals or changes from a continuous to a categorical variable, among others, that might represent SOR or SAR.

We refer to the selected outcome from each CER as the "index outcome" for that CER. We selected this term to differentiate it from the primary outcome of the report, if such was identified, and from the primary outcomes delineated in studies included within each CER.

One of the criteria that we used to select the index outcome for each CER was the number of trials included in the CER that reported the outcome, as described below ("Identification of randomized trials reporting the index outcome within each CER"). Thus the process of identifying an index outcome was an iterative one, where several potential outcomes might have been examined for a CER, and the number of RCTs determined for each of the potential index outcomes.

The index outcome for each of the 15 CERs in our cohort is indicated in Table 3.

Figure 2. Process for identifying an index outcome within each included comparative effectiveness review



Abbreviations: CER, comparative effectiveness review; MA, meta-analysis; RCTs, randomized controlled trials; SOR, selective outcome reporting.

Flow diagram depicting our process for identifying a single, index outcome within each included CER. Bullet points delineate the criteria for selecting an outcome for each CER.

Identification of Randomized Trials Reporting the Index Outcome Within Each Comparative Effectiveness Review

After identifying the index outcome for each of the CERs, we then examined the CER for a list of included RCTs that reported any data on the index outcome. We reviewed all text, tables, and appendices for relevant trials.

We did not review lists of excluded studies, nor did we examine all trials included in the CER. Our approach limited our ability to detect certain types of SOR: we did not examine trials which omitted an outcome completely because those studies would not have been included in the CER for that outcome.

Our approach to identifying trials has important similarities to, and differences from, that described by Kirkham and colleagues in the ORBIT study.[3] Like the ORBIT study, we started with a cohort of systematic reviews, from which we identified included trials. We confined our examination to trials that reported (at least partially), the index outcome. In contrast, ORBIT also examined trials that were included in a review even though they did not report the index outcome. In addition, the ORBIT team examined lists of excluded studies, looking for any potential SOR or SAR. The broader approach to study inclusion taken in the ORBIT study enabled their team to examine types of SOR where the outcome was not reported in the trial (ORBIT classes E, F, G and H; see Table 1). In our study, we did not examine trials that did not report in some way our index outcome.

Similar to the ORBIT study, we did not include trials that were not included in the CER but which may have reported the outcome of interest. In other words, we did not search trial registries or other sources for entire studies that were missing.

Table 3 lists the index outcome that we selected for each of the 15 CERs and the number of RCTs for each index outcome.

Selection of Specific Comparative Effectiveness Reviews for This Pilot Study

In order to have a reasonable number of trials to examine the frequency and predictors of SOR and SAR, we selected the three CERs with the largest number of trials examining each of the three index outcomes.[19, 24, 26] Resource constraints did not permit us to examine all identified CERs in this exploratory work.

We piloted our methods using the published report "Oral Diabetes Medications for Adults with Type 2 Diabetes. An Update" by Bennett and colleagues from the Johns Hopkins EPC.[19] We selected this report because: (1) it was published recently (March, 2011) and therefore likely used current AHRQ methods; (2) diabetes mellitus has a high burden of illness; and (3) this review included a large number of RCTs (n=24 which were published after 2005 and that were listed in a registry) reporting on an outcome that has relatively standardized measurement techniques (A1c).

After developing and piloting our methods on the report on oral diabetes medications, we examined two other CERs.[24, 26] These two CERs contained the largest number of RCTs for the index outcome and we wanted to achieve a reasonable cohort of trials for each of the three CERs. Mortality was chosen for the CER on lipid agents[24] because there was a meta-analysis of that outcome in the report, and we wanted to focus on a patient-important, objective health outcome. We had planned to examine mortality as an outcome of benefit in this report; however RCTs

reporting mortality always considered it an adverse event. We proceeded with an examination of this outcome anyway as our work was well underway when it became clear how mortality was handled in our included studies. For the CER on medications for fracture prevention in women with low bone density,[26] we selected fractures as our "index" outcome because it is a patient-important, objective outcome, and was meta-analyzed in the CER.

A1c, mortality, and fractures thus provided us with a diverse set of outcomes and a cohort of 40 RCTs with which to explore how trial registries might contribute to the assessment of SOR and SAR in CERs.

Identification of Trial Registration for Randomized Controlled Trials Reporting the Index Outcome

In order to identify SOR and SAR for the index outcome in our cohort of CERs and RCTs, we looked for trial registration, in other words, the public listing of an agreed-upon set of information about the design, conduct, and administration of a clinical trial.[18] We focused on registries as a source of information for assessing SOR and SAR in order to develop guidance for the AHRQ EPC program on using registries for this purpose. We were aware of the potential limitations of registries for this purpose, and our goal was to explore the uses and limitations in order to inform guidance on this tool for assessing SOR and SAR.

Through discussion among our workgroup members, we devised an approach for identifying trial registration that optimized the sensitivity of our search (minimized missing registrations), but was still feasible given the available resources. Our approach is outlined in Figure 3, and discussed in detail below.

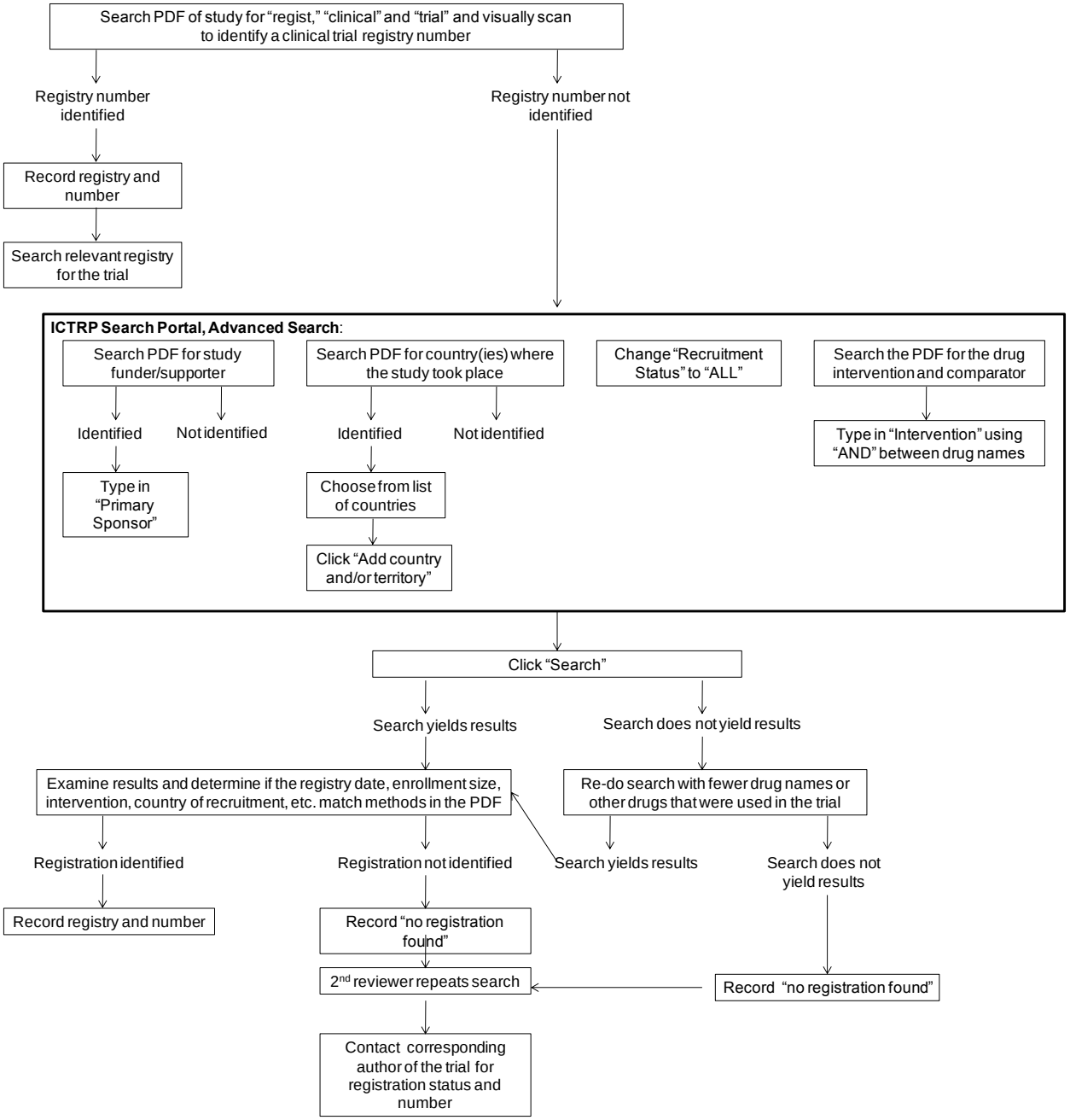
We confined our search for trial registration to studies published from January, 2005 forward because of the effective date of the International Committee of Medical Journal Editors (ICMJE) requirement for registration of published trials (July, 2005). Although ClinicalTrials.gov went online in February, 2000,[35] registration was infrequent initially, and we wanted to focus our search efforts on a cohort of RCTs with a reasonable likelihood of registration. In addition, Zarin and colleagues[36] reported that the quality and completeness of trial records improved since October, 2005.

We searched only the International Clinical Trials Registry Platform (ICTRP) for trial registration and protocols, as this platform encompasses ClinicalTrials.gov and a number of other registries, as noted above, and is updated regularly. If an RCT provided information on trial registration, particularly the National Clinical Trials number for ClinicalTrials.gov, we were able to quickly identify the study within ICTRP. Frequently, however, study publications provided no indication that the study was listed in a registry encompassed by ICTRP. If there was no indication of trial registration in the publication, we then searched for registration using the advanced search tool within ICTRP. We used the study funder, country or countries where subject recruitment took place, and the generic names of the intervention and comparator drugs. If this approach did not reveal the study registration, we then searched on the study drug name only (generic and brand names) as we encountered instances where this less restrictive search identified relevant registrations when the more specific search did not.

This process of searching for trial registrations was repeated for all studies for which we did not initially identify a trial registration by a second, independent searcher. If we still did not identify a study registration, we then contacted the corresponding author indicated in the trial publication via email, and asked if the study was registered, and if so, what the trial registration

number was. For all studies with registration in ICTRP, we looked for results within the relevant registry.

Figure 3. The process for identifying trial registration for each randomized controlled trial



Abbreviations: ICTRP, International Clinical Trials Registry Platform; PDF, portable document format; WHO, World Health Organization.

Data Abstraction

Data from the primary publication, and any companion papers cited in the CER, as well as information from the trial registration were abstracted into a standardized template in Excel (Microsoft Office Excel 2007, Microsoft, Redmond, WA). This template was piloted by study staff at Oregon Health & Science University and by members of the workgroup, using the CER on oral hypoglycemic agents.[19] The template was revised as needed during the pilot phase. Abstracted data were reviewed by a second team member to ensure accuracy. Abstracted data included information from the publication (author affiliations and disclosures, study funder, interventions, relevant outcomes, analytic approach, and analysis set) and registry (dates of registration and recruitment, proposed primary and secondary outcomes and any changes among those, and relevant results [if available]).

Exploration of the Timing of Trial Registration and Changes in Outcomes in the Registry

We compared the time of trial registration to the stages of research, including subject recruitment, study completion, and publication in order to examine how frequently studies were registered after subject recruitment started. In particular, we were interested in the frequency of registration late in the research process, such as at the time of publication.

We also examined the timing of changes in the index outcome within the trial registry. Ideally study outcomes are specified at finalization of the study protocol and prior to subject recruitment and are not subsequently changed. If an outcome did change in any way (e.g., timing of followup, continuous to categorical outcome), a rationale should have been provided in the registry, although this is not currently required in ClinicalTrials.gov.

Identification of Selective Outcome Reporting and Selective Analysis Reporting

To determine the proportion of trials with SOR and/or SAR within a CER, we created a matrix of all the outcomes related to the designated index outcome. This approach was similar to that used by Kirkham and colleagues.[3] Each row in the matrix represented an RCT that had an identifiable trial registration, and each column contained different approaches to presenting and/or measuring the index outcome across the included RCTs and trial registration information. Specifically, each column in the matrix contained outcomes which were:

1. mentioned in the methods section of the publication;
2. reported in the results section of the publication (including outcomes reported in different ways (e.g., categorical [with various definitions of categories] or continuous); or
3. reported in the trial registry (either registration information or results).

This exploration and data display enabled us to depict the universe of potential approaches to specifying and presenting the index outcome for each CER, and to identify outcomes reported in either the trial report or in the trial registry, but not both. We did not seek content expert input to discern other relevant outcomes as we felt that the list of outcomes that we identified in the CER and trial registry adequately represented the universe of potential outcomes.

We used a similar approach to identify SAR. We noted the planned analyses in the methods section of publications and any information in the trial registry, and compared those to the

reported analyses within and across trials. We did not expect that information in trial registries would be as detailed for SAR as for SOR, nonetheless, we sought information on SAR in a similar manner. We did not include information that was provided in the results section of the registry (when available) on how analyses had been performed, as that information was added to the registry with the results, and thus might reflect post hoc decisions and not prespecified approaches.

Since this work was exploratory, we did not develop a specific list of discrepant or altered analyses that we would consider as SAR. Rather, we sought to describe any change in analytic approach between the methods section of the publication or the registry summary information, and results reported in the publication. SAR includes, for example, selective reporting of data on subgroups, presentation of adjusted compared with unadjusted analyses, selection of different analytic sets (e.g., as-treated vs. intention-to-treat), selective approaches to the handling of missing data, choosing to analyze continuously measured variables categorically or changing measurement scales, and choice of cut-point values to define categorical variables.[4] We sought cases where there was a change from the originally specified analytic approach, rather than situations where analyses appropriately tested different models to achieve optimal fit to the data.

In our exploration of SOR and SAR within each included trial, we started with the definition of SOR delineated by Kirkham and colleagues.[3] We soon noted, however, that there were many additional facets of potential reporting biases related to selection of outcomes and analyses that were not encompassed by this definition. We therefore developed a broad set of "judgments" that represent the types of SOR and SAR that we were observing. These judgments were based on the information that was used to determine if SOR or SAR likely exist (Table 4).

Judgment #1 compared the methods and results sections of the trial publication. Judgment #2 compared the "current outcome" listed in the summary information in the registry to the results reported in the publication for that trial. Judgment #3 compared the results reported in the registry to the results reported in the publication. Discrepancies in one of these comparisons constituted an assessment of "yes" for that judgment. Each judgment was made independently by two authors (SLN and HKH), and consensus was achieved through discussion.

Our approach to identifying SOR and SAR focused on discrepancies between the methods and results section of the trial publication, and between information provided in trial registries and the publication. There are several reasons why such discrepancies might occur in addition to SOR and SAR, such as random error or the nonreporting of results for reasons other than the nature and direction of the outcomes. In this pilot study we chose to examine discrepancies, and we did not try and determine the reason for the discrepancies. Nor did we try and determine the risk of bias associated with each discrepancy. Our approach resembled that of the ORBIT study,[3] which focused on discrepancies within the publication (as well as between clinical judgment and the publication).

Table 4. Types of selective outcome reporting and selective analysis reporting classified by the approach used for detection

Judgment #1	Judgment #2	Judgment #3[a]	Comments
Publication methods vs. publication results	**Trial registry summary vs. publication results**	**Trial registry results vs. publication**	
NA	NA	Differences in the numerical values reported in the publication and the registry	---
Change in primary and/or secondary outcome designation for the index outcome between methods and results	Change in primary and/or secondary outcome designation for the index outcome between the registry (current outcome) and the publication	NA	"Yes" if there was a switch in primary vs. secondary designation of the index outcome between the registry and the publication.
Change in the index outcome between methods and results	Change in the index outcome between registry and publication	NA	This includes a change in the specificity of the description of the outcome between the registry and the publication (e.g., outcome listed as "glycemic control" in the registry and "change in A1c" in the publication.
Outcome listed in methods but not reported in results	Outcome listed in the registry but NR in the publication	Outcomes reported in the registry results but not in the publication	Typical SOR per Kirkham and colleagues[3]
Outcome in methods inadequately reported in results	Outcome listed in the registry inadequately reported in publication	Outcomes reported in the registry results but not adequately reported in the publication	Typical SOR per Kirkham and colleagues[3]
Change in followup interval between registry and publication	Change in followup interval between registry and publication	Change in followup interval between registry and publication	Includes any discrepancy between the followup interval specified in the registry and in the publication.
Outcome reported in results but not reported in methods	Outcome reported in publication but not listed in the registry	Outcomes reported in the publication but not in the registry results	Type of SOR that differs from the classes depicted by Kirkham and colleagues[3]
Change in analyses between methods and results	Change in analyses between registry and publication	Change in analyses between registry and publication	---

Table 4. Types of selective outcome reporting and selective analysis reporting classified by the approach used for detection (continued)

Judgment #1	Judgment #2	Judgment #3[a]	Comments
Subgroups reported in results that were not described in the methods section or vice versa	Subgroup specified in the registry but NR in the publication(s) or vice versa	Subgroup reported in the registry but NR in the publication, or vice versa.	---
Summary: judgment #1: SOR and/or SAR are deemed to be present if there are discrepancies in description of the index outcome between the publication methods and results section.	**Summary: judgment #2:** SOR and/or SAR are deemed to be present if there are discrepancies in description for the index outcome between the trial registry summary information and the results section of the publication.	**Summary: judgment #3:** SOR and/or SAR are deemed to be present if there are discrepancies in the results presented for the index outcome between the trial registry results and the results presented in the publication. This includes both the publication of a subset of the results in the registry and the publication of additional results not presented in the registry.	---

Abbreviations: A1c, hemoglobin A1c; NA, not applicable; NR, not reported; SAR, selective analysis reporting; SOR, selective outcome reporting.
[a] Judgment #3 is not applicable if there are no results in the registry or no differences in results between the publication and registry.

In Table 5 we provide a more detailed classification system for SOR and SAR, including the ORBIT categories as well as the type of SOR where outcomes were added that were not prespecified. We also outline the sources for the studies that contribute to the assessment of a given type of SOR and the information that was examined for discrepancies that might represent SOR.

We made the assumption that information in the trial registries was correct as we looked for discrepancies between the registry information and the publication. Issues with the quality of data entry in ClinicalTrials.gov have been raised in the past, although these have focused only on the quality of the entries for the primary outcome measures, particularly their specificity.[36]

Table 5. Types of selective outcome reporting and their relationship to ORBIT classes and to risk of bias

Description of the Outcome	Source of Trials Used To Examine This Type of Outcome	ORBIT Class	Our Assessment	Full, Partial, or No Reporting	Source of Discrepancy Used To Determine if SOR is Present	Risk of Bias[a]	Comment
Outcome fully reported in the results of the publication	Studies included in the review	NA	Judgment #1,2,3 all negative (no SOR)	Full	Between publication methods and publication results, and between other available information (e.g., trial protocol, registry, or clinical judgment) and study publication	None	No SOR (outcome prespecified and fully reported)
	Studies included in the review	NA	One or more of judgments #1,2,3 are positive (outcomes added that were not prespecified)	Full	Between publication methods and publication results, or between other available information (e.g., trial protocol, registry, clinical judgment) and study publication	High	"Data dredging"
Clear that the outcome was measured and analyzed	Studies included in the review	A, B, C	Judgment #1 positive	Partial	Between publication methods and publication results	High or low (per Kirkham)[3]	Outcome is partially reported in the publication
	Studies included in the review	D	Judgment #1 positive	No	Between publication methods and publication results	High (per Kirkham)[3]	The whole outcome is missing or there is a difference in the specific metric for the outcome

Table 5. Types of selective outcome reporting and their relationship to ORBIT classes and to risk of bias (continued)

Description of the Outcome	Source of Trials Used To Examine This Type of Outcome	ORBIT Class	Our Assessment	Full, Partial, or No Reporting	Source of Discrepancy Used To Determine if SOR is Present	Risk of Bias[a]	Comment
Clear that the outcome was measured, unclear whether it was analyzed	Studies included in the review for the index outcome	E, F	Judgment #1 positive	No	Between publication methods and publication results	High or low (per Kirkham)[3]	The whole outcome is missing or there is a difference in the specific metric for the outcome
	Studies included in the review for the index outcome or studies not identified for the review	NA	Judgment #2 positive	No	Between trial registry summary and publication results	High or low	The whole outcome is missing or there is a difference in the specific metric for the outcome
	Studies included in the review for the index outcome or studies not identified for the review	NA	Judgment #3 positive	No	Between trial registry results and publication results	High or low	The whole outcome is missing or there is a difference in the specific metric for the outcome
Unclear whether the outcome was measured	Studies included in the review for other outcomes than the index outcome, or studies not identified for the review	G, H	NA	No	Between clinical judgment and the publication results	High or low (per Kirkham)[3]	The whole outcome is missing or there is a difference in the specific metric for the outcome

Table 5. Types of selective outcome reporting and their relationship to ORBIT classes and to risk of bias (continued)

Description of the Outcome	Source of Trials Used To Examine This Type of Outcome	ORBIT Class	Our Assessment	Full, Partial, or No Reporting	Source of Discrepancy Used To Determine if SOR is Present	Risk of Bias[a]	Comment
Clear that the outcome was measured, unclear whether it was analyzed	Studies included in the review for the index outcome	E, F	Judgment #1 positive	No	Between publication methods and publication results	High or low (per Kirkham)[3]	The whole outcome is missing or there is a difference in the specific metric for the outcome
	Studies included in the review for the index outcome or studies not identified for the review	NA	Judgment #2 positive	No	Between trial registry summary and publication results	High or low	The whole outcome is missing or there is a difference in the specific metric for the outcome
	Studies included in the review for the index outcome or studies not identified for the review	NA	Judgment #3 positive	No	Between trial registry results and publication results	High or low	The whole outcome is missing or there is a difference in the specific metric for the outcome
Unclear whether the outcome was measured	Studies included in the review for other outcomes than the index outcome, or studies not identified for the review	G, H	NA	No	Between clinical judgment and the publication results	High or low (per Kirkham)[3]	The whole outcome is missing or there is a difference in the specific metric for the outcome
Clear that the outcome was not measured	Studies included in the review	I	NA	No	No discrepancy	No risk (per Kirkham)[3]	Bias may have also been introduced at the design phase of the study; this is not captured by SOR because the nonreporting due to study design is not dependent on the results

Table 5. Types of selective outcome reporting and their relationship to ORBIT classes and to risk of bias (continued)

Description of the Outcome	Source of Trials Used To Examine This Type of Outcome	ORBIT Class	Our Assessment	Full, Partial, or No Reporting	Source of Discrepancy Used To Determine if SOR is Present	Risk of Bias[a]	Comment
Clear that the outcome was measured, unclear whether it was analyzed	Studies included in the review for the index outcome	E, F	Judgment #1 positive	No	Between publication methods and publication results	High or low (per Kirkham)[3]	The whole outcome is missing or there is a difference in the specific metric for the outcome
	Studies included in the review for the index outcome or studies not identified for the review	NA	Judgment #2 positive	No	Between trial registry summary and publication results	High or low	The whole outcome is missing or there is a difference in the specific metric for the outcome
	Studies included in the review for the index outcome or studies not identified for the review	NA	Judgment #3 positive	No	Between trial registry results and publication results	High or low	The whole outcome is missing or there is a difference in the specific metric for the outcome
Unclear whether the outcome was measured	Studies included in the review for other outcomes than the index outcome, or studies not identified for the review	G, H	NA	No	Between clinical judgment and the publication results	High or low (per Kirkham)[3]	The whole outcome is missing or there is a difference in the specific metric for the outcome
The study was not published and thus the outcome is not available to the systematic reviewer	Unpublished data (e.g., protocols, trial registries)	NA	NA	No	Published data with no corresponding publication	High	Known as publication bias; not examined in this report

Abbreviations: NA, not applicable; ORBIT, Outcome Reporting Bias in Trials study (Kirkham 2010)[3]; SOR, selective outcome reporting.

[a] Risk of bias is based either on the assessment of Kirkham and colleagues[3]as indicated, or on our subjective assessment of the risk.

Predictors of Selective Outcome Reporting and Selective Analysis Reporting

We explored potential predictors of SOR and SAR, using prespecified, study-level, independent variables. Since this is an exploratory analysis where there is essentially no existing empirical evidence, we selected a variety of diverse predictors based on our hypotheses. First, we predicted that funding and other conflicts of interest on the part of the funder and authors might influence the frequency of reporting biases. We based this prediction on existing data that suggest that financial and other conflicts of interest can affect the data presented and the conclusions in primary studies.[37-40] Second, we sought to examine the possibility that SOR and SAR might vary with the characteristics of the registry and the timing of trial registration. We theorized that the level of detail required in a registry and the timing of registration (ideally before subject recruitment starts) might affect the frequency of SOR. Third, we explored the situation where trial authors changed outcomes in the registry between initial registration and the last update of outcomes in the registry. Since this was not captured in our definitions of SOR and SAR, we hypothesized that such changes might relate to the selection of outcomes based on the results and might therefore correlate SOR and SAR. We therefore examined the following independent variables:

1. Characteristics of the funder and authors
    a. One or more study funders or sponsors manufactured one or more of the intervention drugs (categorical variable, yes/no).
    b. One or more authors were employees of a company making one of the intervention drugs (categorical variable, yes/no; percentage of authors who were employed by the company making one of the intervention drugs (proportions).
    c. Assistance in writing the manuscript was provided by study funder (categorical yes/no). When the source of funding for authorship was unclear or not reported, the response was considered "no."
2. Trial registration
    a. Registry in which study was registered (categorical variable, ClinicalTrials.gov vs. other registry).
    b. Timing of trial registration with respect to first subject recruitment (categorical variable, before or the same month, vs. after or unclear).
    c. Studies with results compared with studies with no results in the registry. Here we only examined judgments #1 (publication methods vs. results) and #2 (registry summary vs. publication results) as the dependent variable. (Note that only ClinicalTrials.gov provided results.)
3. Changes in the index outcome within the registry
    a. Studies that changed the index outcome between "original" and "current" within the trial registry (categorical variable, yes/no). The following were considered "yes" responses:
        i. change in designation between primary and secondary outcome;
        ii. change in specificity (e.g., "glycemic control" was the original outcome and "change in A1c between baseline and week 24" was the current outcome in the trial registry); or

iii. addition of an outcome to the registry (e.g., "percentage of patients who achieved A1c < 7.0 percent at followup" was added to the registry where "change in A1c" was the only original A1c-related outcome).

We examined change in the index outcome (point #3, above) as a predictor of judgments #1, #2, and #3 (our definition of SOR), as changes between the "original" and "current" outcomes listed in trials registries were not part of our definition of SOR. Our criteria for SOR involved only a comparison of the "current" outcome listed in the registry to what was published in the publication (judgment #2: registry summary vs. publication results). Thus a change in outcome from "original" to "current" might be a predictor of SOR as we defined it. In other words, if trial authors made changes to the primary and secondary outcomes listed in the registry, we explored the correlation of such changes with SOR.

The dependent variables for this analysis were:

1. Presence of SOR or SAR detected by examining the publication only ("judgment #1": publication methods vs. publication results).
2. Presence of SOR or SAR detected by comparing registry information to the results reported in the publication ("judgment #2": registry summary vs. publication results).
3. Presence of SOR or SAR detected by comparing the results in the registry (if any) to the results reported in the publication ("judgment #3": registry results vs. publication results).

We combined data from the reports on oral hypoglycemic agents[19] and osteoporosis[26] as we considered that the potential predictors of our judgments on SOR and SAR would likely be similar between the two reports. For the report on lipid modifying agents,[24] the dependent outcomes for all RCTs were judged to have the same type of SOR, precluding statistical analysis to explore predictors.

In order to examine the relationships between the timing of initial study registration, specification of final primary and secondary outcomes, and publication, we explored the chronology of these events in detail by developing a timeline for each RCT for the CER on oral hypoglycemic agents,[19] and osteoporosis[26] based on information from both the study publication and the trial registry.

Assessment of the Reliability and Validity of the ORBIT Classification of Selective Outcome Reporting

We initially proposed examining the inter-rater reliability and validity of the ORBIT classification system, in addition to the feasibility of its implementation. For our assessment of inter-rater reliability, each index outcome within each included RCT was classified by two independent assessors as fully reported, partially reported, or not reported, using the ORBIT categories A through F.[3] The assessors (SLN and HKH), both coauthors on this report, have formal training in epidemiology, and are experienced in systematic review methods including the assessment of risk of bias in primary research studies. After documentation of the two independent assessments, the two assessors achieved consensus through discussion. The assessors initially piloted the ORBIT assessments by independently rating four included RCTs, followed by discussion and consensus on the ORBIT classification. For the RCTs that were not

part of the training exercise, we planned to calculate a kappa statistic as well as percentage agreement between assessors for each of the nine ORBIT categories and overall.

We planned to assess the validity of the ORBIT classification system by comparing the ORBIT classification of SOR using the trial publication to the ORBIT classification achieved using additional information from the trial registry. We planned to examine the percentage agreement in the classification between these two approaches.

We examined one aspect of the feasibility of using ORBIT by recording the time in minutes it took each assessor to complete the ORBIT classification using the trial publication and registry information, and then the time it took to examine trial results in the registry (if any) and to make any further assessments of ORBIT.

Data Syntheses and Analyses

This was an exploratory study, and so we selected a convenience sample of CERs and RCTs to examine, and no sample size calculations were performed. The three CERs that we examined involved different types of outcomes, so we described each of the three cohorts of RCTs separately.

Descriptive statistics were used to present our findings on SOR and SAR among the RCTs included in this study. In addition, we examined the association between potential predictors and the presence or absence of SOR or SAR as indicated by judgments #1, #2, and #3 as outlined above. Due to the small sample size, the association between the presence of SOR and SAR and study-level characteristics was explored using Fisher's exact test and exact logistic regression. Since there was no significant difference between SOR and SAR in the two CERs (diabetes and fracture prevention), the analysis was first conducted by combining data from the diabetes and fracture studies to assess overall association. Then the analysis was conducted by examining RCTs in the two CERs separately to look at association for each individual CER.

Results

For the three CERs in this exploratory study, we identified a total of 40 RCTs that fulfilled our inclusion criteria: trials published in 2005 or later that reported the index outcome and for which we were able to identify trial registration. Twenty-four of these trials were included in the CER on oral hypoglycemic agents,[19] 14 on lipid-modifying agents[24] and 12 on treatments to prevent fractures.[26]

Table 6 summarizes the characteristics of the RCTs for the index outcomes of the three CERs. Registered trials variably reported trial registration numbers in the publications (28.6, 66.7, and 75.0 percent for the CERs on lipid modifying agents, fracture prevention drugs, and oral hypoglycemic agents, respectively). The majority of registered studies were listed in ClinicalTrials.gov, although this varied across the three CERs. The percentage of trials with results posted in the registry also varied (8.3, 14.3, and 62.5 percent for fracture prevention drugs, lipid modifying agents, and oral hypoglycemic agents, respectively). Of trials published from September, 2008 forward (when results reporting was mandated for some trials within ClinicalTrials.gov), 81 percent of the hypoglycemic agent trials and 100 percent of the lipid modifying agent trials were registered.

We contacted the authors of trials reporting the index outcome when there was no indication of study registration in the publication. In total we attempted to contact 79 corresponding authors by email; 12 email addresses were no longer valid. A total of 26 authors responded (39 percent of authors with valid addresses), and of those, 23 authors indicated that their trials were not registered and three authors provided trial registry numbers.

Table 6. Characteristics of trials that reported the index outcomes

CER	Index Outcome	Total Unique RCTs in the CER (n)	RCTs Published 2005 to January 2012 (n)	Registered Trials (n [% 2005-January 4, 2012;[a] % all years[b]])	Trials with Registry Number Reported in the Publication n (%)	Trials Included in ClinicalTrials.gov (n [% of Trials Registered])	Trials Included in ISRCTN Registry (n [% of Trials Registered])	Trials Listed in Registries Not in ICTRP (n [% of Trials Registered])	Trials with Results Reported in the Registry (n [% of Trials Registered; % Published September 2008 to January 2012])
Oral Diabetes Medications for Adults With Type 2 Diabetes[19]	A1c	98	59	24 (40.7; 24.5)	18/24 (75)	23 (95.8)	0 (0.0)	1 (4.2) UMIN Clinical Trials Registry[c]	15 (62.5; 81.3)
Comparative Effectiveness of Treatments To Prevent Fractures in Men and Women With Low Bone Density or Osteoporosis[26]	Fracture reduction, including both vertebral and extremity, hip, and other sites	177	29	12 (41.4; 6.8)	8/12 (66.7)	8 (66.7)	2 (16.7)	2 (16.7) Australia New Zealand Clinical Trials Registry	1 (8.3; NA)
Comparative Effectiveness of Lipid-Modifying Agents[24]	All-cause mortality	24	19	14 (73.7; 58.3)	4/14 (28.6)	11 (78.6)	2 (14.3)	1 (7.1%) Cochrane Renal Group[c]	2 (14.3; 100.0)

Abbreviations: A1c, hemoglobin A1c; CER, comparative effectiveness review; ICTRP, International Clinical Trials Registry Platform; ISRCTN, International Standard Randomized Controlled Trial Number; n, number; RCTs, randomized controlled trial

[a] Percentage of trials reporting the index outcome published in or after 2005 that were registered.

[b] Percentage of all trials reporting the index outcome that were registered.

[c] Two publications indicated trial registration in registries not contained within the International Clinical Trials Registry Platform.

Frequency and Characterization of Selective Outcome Reporting and Selective Analysis Reporting

Tables 7–10 summarize our findings on SOR and SAR, including our three consensus judgments on the presence of SOR and SAR as defined previously in Table 3. Additional details on the studies included in our analyses are provided in Appendix tables B1 to B16.

Hemoglobin A1c Outcomes

The cohort of trials reporting A1c (n=24) (Tables 7–9), reported this outcome in a variety of ways, both as a continuous and as a categorical variable (Table 7). In none of these trials did the designation of A1c as a primary or secondary outcome change within the publication and the outcomes listed in the methods section were always reported in the results. In two studies (8 percent) the outcome was not fully reported in the results (judgment #1: publication methods vs. results). In three studies (12.5 percent), however, A1c outcomes were reported in the results sections that were not mentioned in the methods section. In addition, in five studies (21 percent) subgroup data were presented on A1c that were not specified in the methods section of the publication, and in one other study the analytic approach changed between the methods and results section of the publication.

Thus in 46 percent (n=11) of the trials reporting A1c we judged SOR and/or SAR to be present based only on the publication (judgment #1: publication methods vs. results). In seven of those 11 studies (64 percent) we were able to assign an ORBIT class, but in the remaining four we could not. These latter situations were where outcomes were presented in the results that were not mentioned in the methods section.

When the A1c trial publications were examined along with information in the registry summary (but not the registry results) (judgment #2: registry summary vs. publication results), SOR or SAR were assessed as present in 20 studies (83 percent) (Table 9). In none of these trials could we designate an ORBIT class, as all entailed the addition of results to the trial publication that were not specified in the registry. Of the 16 trials with results in the registry, two were assessed as having SOR or SAR (judgment #3: registry results vs. publication results), and ORBIT classes did not apply as these two studies had results added to the publication that were not in the registry.

We report the frequency of prespecified potential predictors of SOR and SAR for A1c trials in Table 10. Ninety-two percent of studies had one or more authors who were an employee of the company sponsoring the study and making one or more of the intervention drugs. Almost half of studies (46 percent) received assistance from the study funder in authoring the publication. Also in 46 percent of studies, subject recruitment started before the trial was registered.

Fracture Outcomes

The cohort of trials reporting the outcome of fractures (n=12) is summarized in Tables 8–9. These studies were much better reported than the A1c cohort, with only four studies (33 percent) reporting outcomes in the results section that were not mentioned in the methods section of the publication (judgment #1: publication methods vs. results). Likewise, there was much better agreement between trial registry outcome information and the reported results, although judgment #2 (registry summary vs. publication results) was still positive in 58 percent of studies. ORBIT class did not apply in any of these trials where SOR or SAR was judged to be present. Only one study of the eight registered in ClinicalTrials.gov had results reported in the registry

and the results differed between the registry and the publication (judgment #3). Industry employment of the study authors (33 percent) was less common than for the A1c trials (92 percent).

Mortality Outcomes

Among the cohort of trials that reported mortality outcomes with lipid agents (n=14), this outcome was always reported as an adverse events (Tables 6 and 11). Eight trials mentioned safety or adverse events in the methods section of the publication, but only one explicitly mentioned death as an outcome in the methods section (judgment #1: publication methods vs. results). None of these studies mentioned mortality as a primary or secondary outcome in the registry summary information (judgment #2: registry summary vs. publication results), and neither of the two studies reporting results in the registry mentioned mortality (judgment #3: registry results vs. publication results). As a result of the homogeneity of reporting across these studies and the fact that none of these studies prespecified mortality as a primary or secondary outcome of either benefit or harm, we did not explore this cohort of studies further.

Table 7. Selective outcome reporting and selective analysis reporting for hemoglobin A1c (n=24) based on the publication only (judgment #1)

	Outcomes in Trials Related to A1c					Discrepancies Between Publication Methods and Results							
No/Yes	Change in A1c From Baseline	Subjects With A1c <7.0%	Subjects With A1c <6.5%	Subjects with a Decrease in A1c > Specific Value	Other A1c Outcome	Change in Primary and/or Secondary Outcome	Outcome in Methods Not Reported in Results	Outcome in Methods Not Fully Reported in Results	Outcome in Results Not Listed in Methods	Change in Analyses Between Methods and Results	Subgroup Results Not Described in Methods	SOR/ SAR Present: Judgment #1	ORBIT Class
No (%)	0 (0)	8 (33)	12 (50)	23 (96)	22 (92)	24 (100)	24 (100)	22 (92)	21 (88)	23 (96)	19 (79)	13 (54)	No SOR/ SAR: 13
Yes (%)	24 (100)	16 (67)	12 (50)	1 (4)	0 (0)	0 (0)	0 (0)	2 (8)	3 (13)	1 (4)	5 (21)	11 (46)	A: 2;[a] B:2; D: 3; E: 1; NA: 4

Abbreviations: NA, not applicable; ORBIT, Outcome Reporting Bias in Trials; SAR, selective analysis reporting; SOR, selective outcome reporting.
A, B, D, E are ORBIT classes (see Table 1).
[a] One of the 11 publications with SOR/SAR was assigned two ORBIT classes.
Data are number of trials (percentage of all trials where n=24 [A1c] and n=12 [fractures]), unless otherwise specified.

Table 8. Selective outcome reporting and selective analysis reporting for fractures (n=12) based on the publication only (judgment #1)

	Types of Fractures Outcomes						Discrepancies Between Publication Methods and Results							
No/Yes	Total	Hip	Vertebral	Non Vertebral	Wrist	Other	Change in Primary and/or Secondary Outcome	Outcome in Methods Not Reported in Results	Outcome in Methods Not Fully Reported in Results	Outcome in Results Not Listed in Methods	Change in Analyses Between Methods and Results	Subgroup Results Not Described in Methods	SOR/ SAR Present: Judgment #1	ORBIT Class
No (%)	1 (8)	7 (58)	6 (50)	7 (58)	8 (67)	7 (58)	12 (100)	12 (100)	12 (100)	8 (67)	12 (100)	12 (100)	8 (67)	No SOR/ SAR: 8
Yes (%)	11 (92)	5 (42)	6 (50)	5 (42)	4 (33)	5 (42)	0 (0)	0 (0)	0 (0)	4 (33)	0 (0)	0 (0)	4 (33)	ORBIT NA: 4

Abbreviations: NA, not applicable; ORBIT, Outcome Reporting Bias in Trials; SAR, selective analysis reporting; SOR, selective outcome reporting.

A, B, D, E are ORBIT classes (see Table 1).

[a] One of the 11 publications with SOR/SAR was assigned two ORBIT classes.

Data are number of trials (percentage of all trials where n=24 [A1c] and n=12 [fractures]), unless otherwise specified.

Table 9. Selective outcome reporting and selective analysis reporting for A1c and fracture outcomes based on the publication and on the registry (judgments #2 and #3)

Outcomes (Number of Studies)	No/Yes	Registry Without Consideration of Registry Results							With Registry Results	
		Change in Outcome Between Registry and Publication	Change in Status of Outcome[a] Between the Registry and the Publication	Change in Followup Interval Between Registry and Publication	Outcome Not Reported in Publication but Listed in Registry	Outcome Reported in Publication but Not Reported in Registry	SOR/SAR Present: Judgment #2 (Registry Summary vs. Publication Results)	ORBIT Class (Judgment #2)	SOR/SAR Present: Judgment #3 (Registry Results vs. Publication Results)	ORBIT Class (Judgment #3)
A1c outcomes (n=24)	No (%)	23 (96)	24 (100)	21 (88)	24 (100)	4 (17)	4 (17)	No SOR/SAR: 4	(No results: 8)	No SOR/SAR: 6 SOR/SAR NA: 8
	Yes (%)	1 (4)	0 (0)	2 (8)	0 (0)	20 (83)	20 (83)	ORBIT NA: 20	2/16 (13)	ORBIT NA: 2
Fracture outcomes (n=12)	No (%)	12 (100)	12 (100)	9 (75)	12 (100)	5 (42)	5 (42)	No SOR/SAR: 5	(No results: 11)	SOR/SAR NA: 11
	Yes (%)	0 (0)	0 (0)	3 (25)	0 (0)	7 (58)	7 (58)	ORBIT NA: 7	1/12 (8.3)	ORBIT NA: 1

Abbreviations: NA, not applicable; ORBIT, Outcome Reporting Bias in Trials; SAR, selective analysis reporting; SOR, selective outcome reporting.

[a] Change in designation of the index outcome as primary or secondary.

Data are number of trials (percentage of all trials where n=24 [A1c] and n=12 [fractures]), unless otherwise specified.

Table 10. Potential predictors of selective outcome reporting and selective analysis reporting for A1c and fracture outcomes

Outcomes	No/Yes	Author With Industry Affiliation	Study Drug Made by Study Sponsor[a]	Assistance Authoring the Publication by Study Funder	Trial Registered in ClinicalTrials.gov	First Subject Recruited Before Trial Registered	Registry Reports Results	Change in the Index Outcome Between Original and Current Outcome Listed in the Registry
A1c outcomes (n=24 trials)	No (%)	2 (8)	0 (0)	13 (54)	1 (4)	13 (54)	8 (33)	19 (79)
	Yes (%)	22 (92)	23 (96)	11 (46)	23 (96)	11 (46)	16 (66)	5 (21)
Fracture outcomes (n=12 trials)	No (%)	8 (67)	6 (50)	10 (83)	4 (33)	11 (92)	11 (92)	11 (92)
	Yes (%)	4 (33)	6 (50)	2 (17)	8 (67)	0 (0)	1 (8)	1 (8)

Abbreviations: SAR, selective analysis reporting; SOR, selective outcome reporting.
[a] The trial sponsor was not reported in one of the A1c trials.
Data are number of trials (percentage of all trials where n=24 [A1c] and n=12 [fractures]) unless otherwise specified.

Table 11. Reporting of mortality in studies of lipid-modifying agents (n=14 trials)

No/Yes	Trials With Adverse Events Mentioned in the Methods Section of the Publication	Trials With Mortality Mentioned in the Methods Section	Trials With Adverse Events Mentioned in the Registry Summary	Trials With Results in the Registry	Trials With Mortality Reported in the Registry
No (%)	8 (57)	13 (93)	14 (100)	12 (86)	14 (100)
Yes (%)	6 (43)	1 (7)	0 (0)	2 (14)	0 (0)

Data are number of trials (percentage of all trials where n=14).

Timing of Study Registration

Trial registration occurred at various times during the course of included trials (Table 12). A minority of studies was registered before or in the same month that subject recruitment started (46 percent, 29 percent, and 0 percent in the CERs on oral hypoglycemic agents, lipid-modifying agents, and fracture prevention drugs, respectively). Studies that commenced before 2005 (the year ICMJE recommendations were implemented) reasonably could have been registered later in the course of the study. However, ClinicalTrials.gov was launched in 2000 so that registration could have occurred prior to commencement of subject recruitment for virtually all studies in our cohort which were all published in 2005 or later.

Table 12. Timing of study registration

Time of Study Registration	Oral Diabetes Medications for Adults With Type 2 Diabetes[19] (n=24)	Comparative Effectiveness of Lipid-Modifying Agents[24] (n=14)	Fractures in Men and Women With Low Bone Density or Osteoporosis[26] (n=12)
Before recruitment started or in the same month that recruitment started	11 (45.8%)	4 (28.6%)	0 (0.0%)
Before primary completion date, primary completion date for recruitment, or date recruitment completed (from publication)	16 (66.7%)	4 (28.6%)	5 (41.7%)
Before study completion date	18 (75.0%)	7 (50.0%)	7 (58.3%)
Before publication submitted	19 (79.2%)	10 (71.4%)	8 (66.7%)
Before publication accepted	21 (87.5%)	10 (71.4%)	9 (75.0%)
Before article first published online	21 (87.5%)	10 (71.4%)	10 (83.3%)
Before publication printed	21 (87.5%)	11 (78.6%)	10 (83.3%)
Before date of our review (Jan. 4, 2012)	24 (100.0%)	14 (100.0%)	12 (100.0%)

Numbers in each row are cumulative.

Changes in the Index Outcome in the Registry

We examined the proportion of trials with a change in the index outcome between the "original" outcome and the "current outcome" in the trial summary in the registry, and the timing of those changes (Table 13). Forty-two percent of trials in the CER on oral hypoglycemic agents had a change in some aspect of A1c measurement, whereas 16.7 percent of trials in the CER on fracture prevention had a change in the index outcome in the registry. The outcome of mortality was never mentioned in the trial summaries for the CER on lipid management, so there was no documentation of any changes.

The timing of changes in the index outcome between the original and current outcome in the registry was variable (Table 13), and occurred late in the research and publication process. In seven of 24 RCTs examining A1c and in two of 12 RCTs reporting fractures, the change occurred after the date of publication of the trial.

We also examined whether our index outcome changed with respect to the posting of results in ClinicalTrials.gov, by reviewing the "History of Changes" in the registry. For the CER on oral hypoglycemic agents,[19] of the 10 RCTs with changes in the specification of A1c, this outcome changed twice on the date the results were posted to the registry, and six times after the date the results were posted. (In two RCTs the results were not posted.) For the two trials where the

outcome of fractures changed in the registry, one change occurred after the date the results were posted, and in the other trial no results were posted.

Table 13. Timing of changes in the index outcome in the registry

Time of Change in Index Outcome	Oral Diabetes Medications for Adults With Type 2 Diabetes[19] (n=24)	Fractures in Men and Women With Low Bone Density or Osteoporosis[26] (n=12)
Before recruitment started or in the same month that recruitment started	0	0
Before the primary completion date or the primary completion date for recruitment or before date that recruitment was completed (from trial publication)	0	0
Before study completion date	0	0
Before publication submitted to publishing journal	1	0
Before publication accepted	2	0
Before article first published online	2	0
Before publication printed	3	0
Before date of our review (Jan. 4, 2012)	10	2
Total number of trials with a change in the index outcome (%)	10/24 (41.7%)	2/12 (16.7%)

Data are the cumulateive number of trials with a change in the index outcome for each time interval.

We illustrate the chronology of trial registration and changes in outcomes with three examples (Figures 4, 5, and 6). Dates above the horizontal line refer to dates from the publication; dates below the line are from the trial registry. Figure 4 provides an example of a reasonable study-registration chronology, wherein the study was registered at the time it was started and results were posted. Ideally, however, results would have been posted earlier, at the time of publication. Figures 5 and 6 illustrate suboptimal chronologies. In Figure 5, the trial was registered after the study was submitted for publication, and the results submitted 2 years later. In Figure 6 the study was also registered after publication and results were never reported.

Figure 4. Example of study and registry chronology (from Raz et al., 2008)[41]

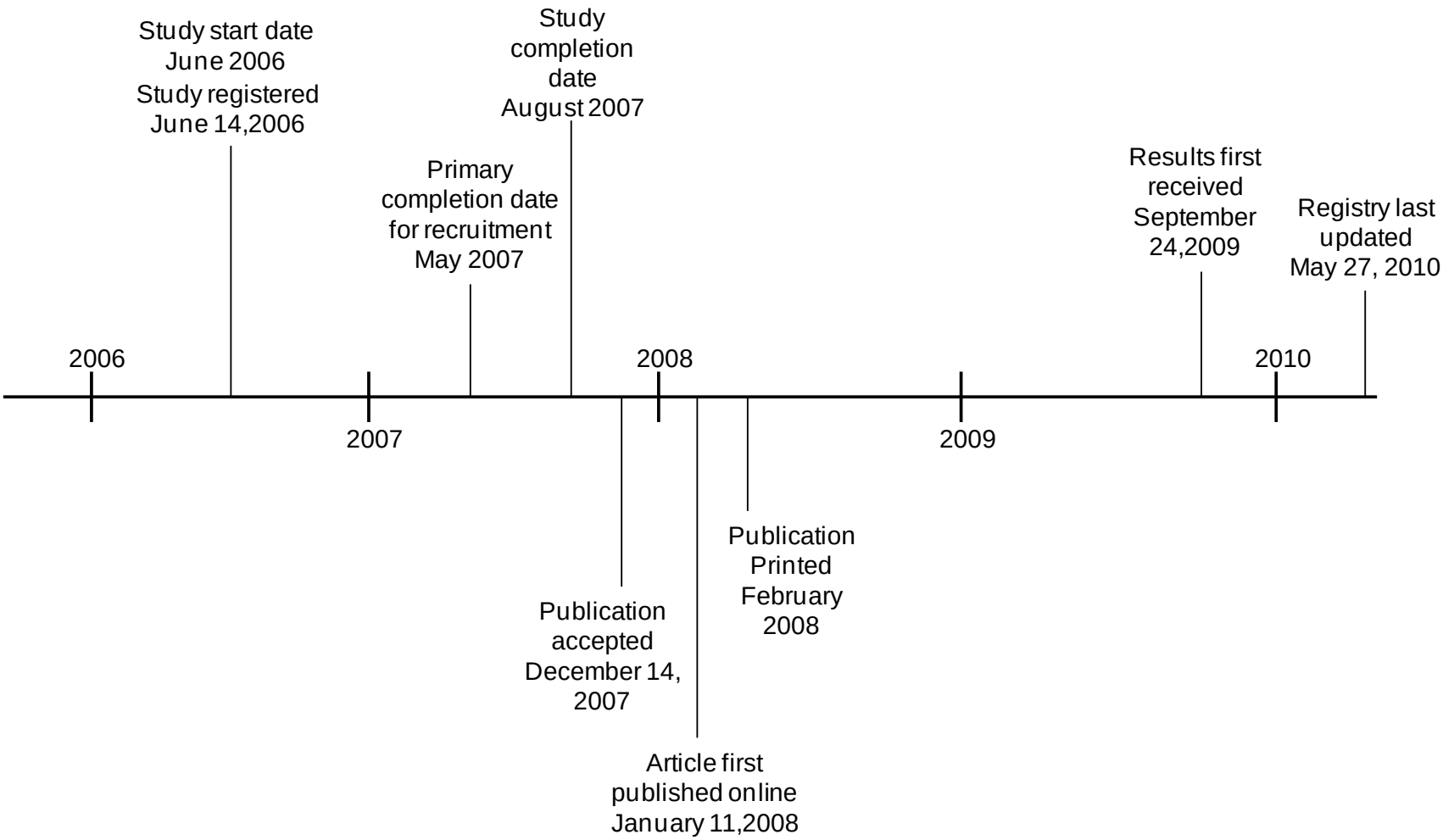


Figure 5. Example of study and registry chronology (from Scott et al., 2008)[42]

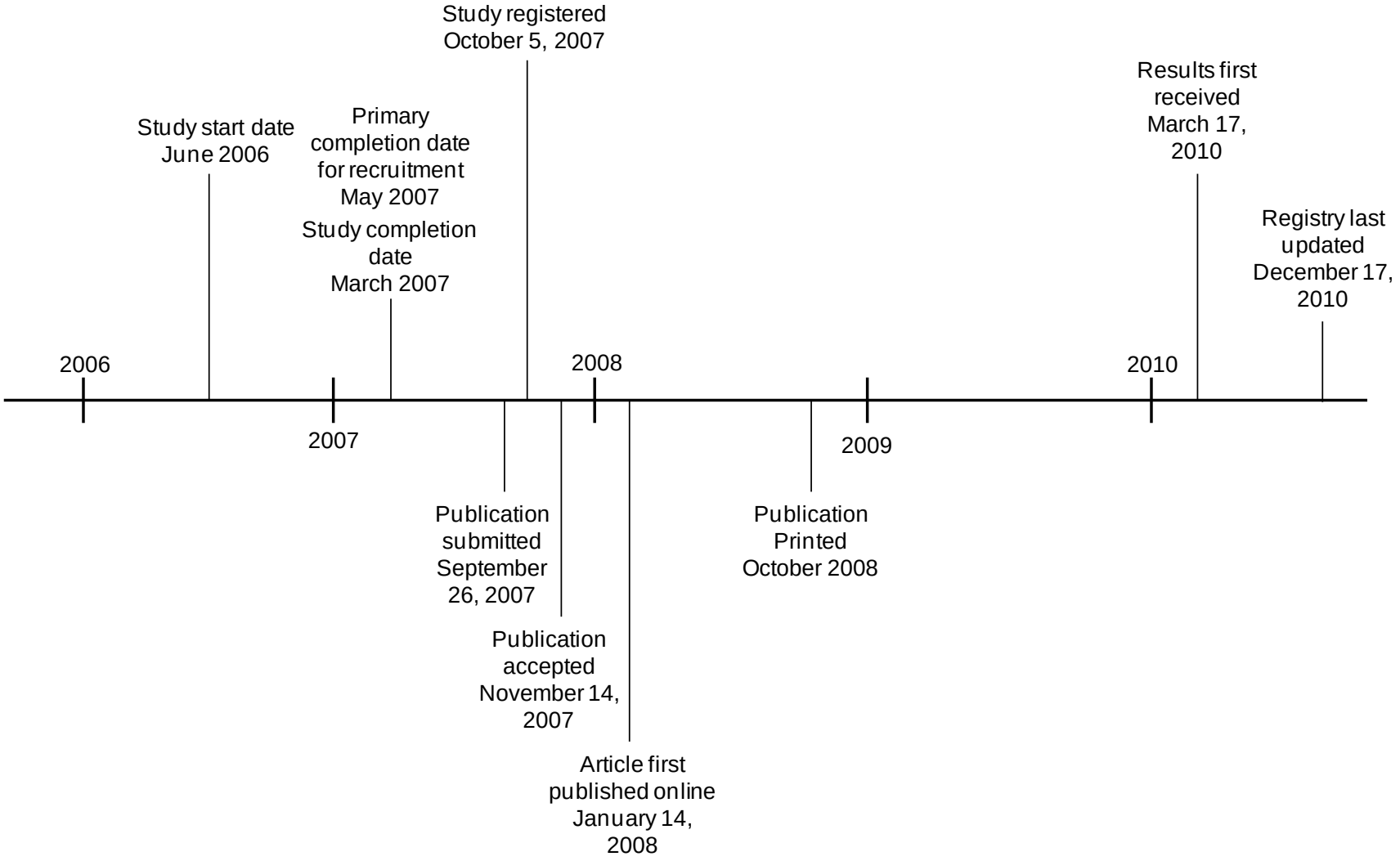
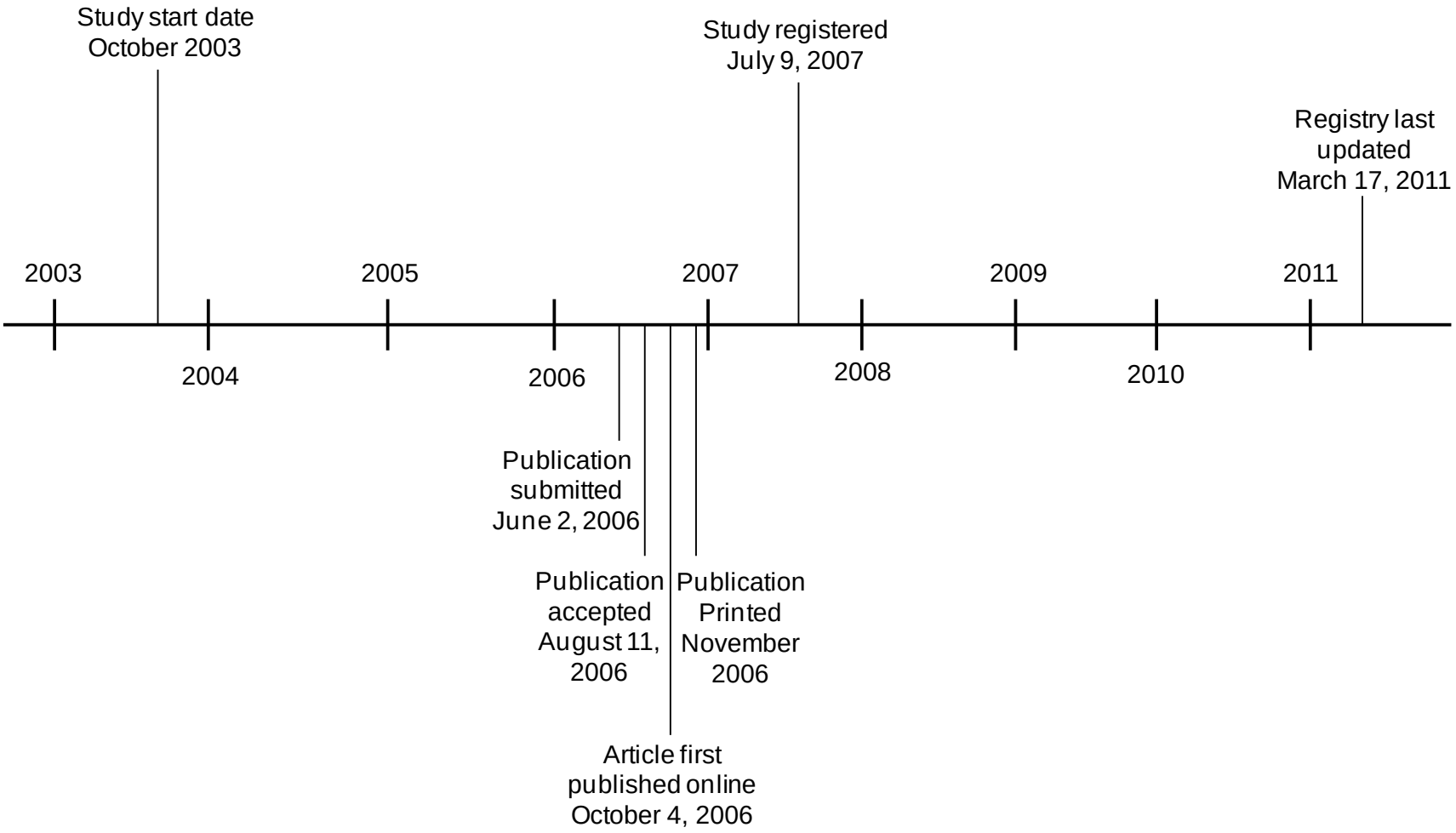


Figure 6. Example of study and registry chronology (from Rosenstock et al., 2006)[43]



Efforts Taken To Identify Selective Outcome Reporting and Selective Analysis Reporting

The time it took the two assessors to examine the trial registry for SOR and SAR is reported in Table 14. There was considerable variation across studies: examination of the summary information in the registry took as few as 3 minutes and as long as 25 (median 10, interquartile range [IQR] 6.5). Examination of trials with results in the registry took the two assessors 21 and 23 minutes on average for the studies examining A1c. We did not collect data for the report on lipid agents.

Table 14. Time to complete the review of ClinicalTrials.gov

Report	Mean Time Without Examination of Results (n, [range]; median; IQR)		Mean Time with Examination of Results (n, ([range]; median; IQR)	
	Reviewer 1	Reviewer 2	Reviewer 1	Reviewer 2
Diabetes (n=24)	13 (5 to 23); 13; 5.5 n=23	10 (3 to 25); 9; 6 n=18	23 (16 to 38); 21; 10 n=14	21 (9 to 34); 19; 8.5 n=12
Lipids (n=14)	NA	NA	NA	NA
Osteoporosis (n=12)	10 (4 to 15); 9.5; 4 n=11	7 (3 to 15); 6; 3 n=12	17 (17 to 17); 17; 0 n=1[a]	19 (19 to 19); 19; 0 n=1[a]

Abbreviations: NA, not applicable.
[a] The osteoporosis review only had one trial with results reported.
Time in minutes.

Predictors of Selective Outcome Reporting

We explored the relationships between various prespecified characteristics of study authors, trial registration, and changes in the index outcome within the registry, and our assessment of SOR and SAR (judgments #1, #2, and #3 for the CERs on oral hypoglycemic agents and drugs for prevention of fractures). We analyzed both CERs separately and combined, and found no significant relationships (all p values >0.05), although our small sample size limited power to detect significant differences (Tables 15 and 16). There was not sufficient variation in the judgments of SOR and SAR for the CER on lipid-modifying agents and mortality to perform predictor analyses for that report.

Table 15. Oral hypoglycemic agents: Relationship between potential predictors and the judgment on selective outcome reporting and selective analysis reporting

Predictor	Total N	Judgment #1	p value	Total N	Judgment #2	p value	Total N	Judgment #3	p value
One or more authors with industry affiliation[a]	N (n=2) Y (n=22)	0 11	0.482	N (n=2) Y (n=22)	2 18	1.000	N (n=0) Y (n=16)	NA	NA
Intervention drug made by study sponsor[a]	U (n=1) Y (n=23)	1 10	0.458	U (n=1) Y (n=23)	1 19	1.000	U (n=1) Y (n=15)	0 2	1.000
Assistance authoring the publication by study sponsor[a]	N (n=13) Y (n=11)	6 5	1.000	N (n=13) Y (n=11)	10 10	0.596	N (n=8) Y (n=8)	2 0	0.467
Study registered in ClinicalTrials.gov[a]	N (n=1) Y (n=23)	0 11	1.000	N (n=1) Y (n=23)	1 19	1.000	N (n=0) Y (n=16)	NA	NA
First subject recruitment occurred before (or during same month) study started[a]	N (n=13) Y (n=11)	5 6	0.682	N (n=13) Y (n=11)	11 9	1.000	N (n=6) Y (n=10)	1 1	1.000
Registry reported the results[a]	N (n=9) Y (n=15)	3 8	0.423	N (n=9) Y (n=15)	7 13	0.615	N (n=1) Y (n=15)	0 2	1.000
Change in A1c measure between original and current registry outcomes[a]	N (n=19) Y (n=5)	8 3	0.63	N (n=19) Y (n=5)	15 5	0.544	N (n=12) Y (n=4)	0 2	0.050
Number of authors with industry affiliation[b]	OR (95% CI)	0.839 0.505-1.322	0.5095	OR (95% CI)	1.028 0.561-1.976	1.000	OR (95% CI)	1.279 0.500-3.056	0.6167
Percent of authors with industry affiliation[b]	OR (95% CI)	0.979 0.948-1.011	0.203	OR (95% CI)	0.995 0.954-1.037	0.802	OR (95% CI)	0.990 0.929-1.056	0.768

Abbreviations: Y, Yes; N, No; U, Unsure; A1c, hemoglobin A1c; CI, confidence interval; NA, not applicable; OR, odds ratio.

[a] Fisher's Exact test

[b] Logistic regression

Studies can have more than one judgment as "yes", thus the rows can add up to more than the total number of trials.

Judgment #1: Publication methods compared with the publication results.

Judgment #2: Trial registry summary information compared with the publication results.

Judgment #3: Trial registry results compared with the publication results.

Table 16. Drugs for the prevention of fractures: Relationship between potential predictors and the judgment on selective outcome reporting and selective analysis reporting

Predictor	Total N	Judgment #1	p value	Total N	Judgment #2	p value
One or more authors with industry affiliation [a]	N (n=8) Y (n=4)	2 2	0.547	N (n=8) Y (n=4)	5 2	1.000
Intervention drug made by study sponsor [a]	N (n=6) Y (n=6)	1 3	0.545	N (n=6) Y (n=6)	4 3	1.000
Assistance authoring the publication by study sponsor [a]	N (n=10) Y (n=2)	3 1	1.000	N (n=10) Y (n=2)	6 1	1.000
Study registered in ClinicalTrials.gov [a]	N (n=4) Y (n=8)	0 4	0.208	N (n=4) Y (n=8)	2 5	1.000
First subject recruitment occurred before (or during same month) study started [a]	N (n=11) Y (n=1)	4 0	1.000	N (n=11) Y (n=1)	6 1	1.000
Registry reported the results [a]	N (n=11) Y (n=1)	3 1	0.333	N (n=11) Y (n=1)	6 1	1.000
Number authors with industry affiliation [b]	OR (95% CI)	1.337 0.752-2.722	0.3879	OR (95% CI)	0.960 0.560-1.677	0.975
Percent authors with industry affiliation [b]	OR (95% CI)	1.075 0.980-1.178	0.126	OR (95% CI)	0.040 0.929-1.086	0.909

Abbreviations: Y, Yes; N, No; U, Unsure; CI, confidence interval; OR, odds ratio.
[a] Fisher's Exact.
[b] Logistic regression.
Judgment #1: Publication methods compared with the publication results.
Judgment #2: Trial registry summary information compared with the publication results.
Judgment #3: Trial registry results compared with the publication results.

Reliability and Validity of ORBIT

We were not able to examine the inter-rater reliability of the ORBIT classification system. for two reasons. First, because of the design of our study, very few trials were assessed as having an ORBIT class, and, by design, ORBIT classes G and H were not identified. We could potentially have examined the reliability of ORBIT classes A through D, but did not have sufficient number of studies with those classes to perform that assessment.

Second, the two assessors found that more trials than the initially proposed pilot of four trials were required to understand the types of SOR and SAR. Each additional trial appeared somewhat different in its presentation of SOR and SAR, and there were a multitude of nuances. The two assessors therefore altered their categorization of SOR and SAR (judgments #1, #2, and #3) during their consensus process. We did not feel that an assessment of inter-rater reliability would be particularly useful with such a process.

Our proposed assessment of the validity of ORBIT proved infeasible, in large part due to our study design. Because we had restricted our cohort of trials to those that reported the index outcome, that outcome (defined broadly, e.g., A1c) had to be reported in our cohort of RCTs and ORBIT classes G and H were thus not identified. Only these two ORBIT classes are applicable to a validation study with comparison of the outcomes that should have been reported (based on clinical judgment) to information contained in sources outside of the study publication (i.e., to registry information in our study).

Discussion

Our work identifying and characterizing SOR and SAR in a small cohort of trials and using information from the study publication and the trial registry suggests that outcomes are frequently added to the results in publications that were not listed in the publication's methods section or specified in trial registries. Types of SOR that correspond to those in the ORBIT classification system A through F were uncommon in our cohort. Trial registries indicated frequent changes in the outcomes listed in the study summary, with many of those changes occurring late in the research process. In addition, there was often inadequate specification of the initial outcomes, making it difficult to determine if there were significant differences in outcomes between the registry and publication. Neither the study publication nor the registry provided useful information for identifying SAR. In addition, exploration of trial registries was time consuming. The identification of trial registry numbers took multiple steps and trial registries did not have an optimal interface for exploring SOR or SAR.

Our work provides valuable experiences and lessons upon which to start building operational guidance for exploring SOR and SAR in trials and for incorporating those findings into systematic reviews.

Strengths and Limitations of This Study

We explored SOR and SAR in a real-world cohort of RCTs included in AHRQ CERs. The challenges encountered and lessons learned contribute to knowledge directly applicable to future CERs supported by AHRQ and other sponsors and authors. We are not aware of published literature exploring the use of ORBIT, nor other attempts to examine the reliability and validity of this classification system. We purport that our suggestions for revisions of existing tools and our recommendations for future research may be applicable to work on reporting biases in biomedical studies, well beyond the cohort of trials and CERs examined in this report.

There are a number of limitations to our study. First, we examined only three CERs and 40 trials. Thus our findings may not be applicable to systematic reviews of other types of interventions and outcomes, and our findings on SOR and SAR at the trial level may not be generalizable to a wider variety of trials. Registration of studies of designs other than randomized trials is infrequent (observational studies constituted 15 percent of studies in ClinicalTrials.gov in January, 2007[36]) and the prevalence of SOR and SAR may exceed that for trials.[44, 45] Thus, our findings are unlikely to be applicable to study designs other than RCTs. In addition, our small sample size of RCTs limited our power to detect significant predictors of SOR.

Second, we faced limitations inherent in trial registration: lack of registration of all trials after 2005 and infrequent registration prior to that year. Thus we could not examine complete cohorts of trials included in meta-analyses within CERs.

Third, studies listed in trials registries may not be representative of all trials published. Perhaps registered trials differ in important reporting and quality characteristics from trials that are not registered. If registered trials are of higher quality, they might have lower rates of SOR and SAR, and thus our cohort may underestimate the frequency of SOR and SAR across trials included in a CER, both registered and unregistered.

Fourth, information in various trial registries differed. Most importantly, more information was provided in ClinicalTrials.gov than in other registries for our cohort of trials. Thus all registered trials do not contain the same information and level of detail. For example, as far as

we are aware, only ClinicalTrials.gov tracks all changes made to the registration information. Thus comparing and synthesizing the frequency of SOR and SAR across different registries is problematic.

Fifth, during the consensus process between the two assessors of SOR/SAR, our thinking and definitions of SOR/SAR evolved as each assessment brought nuances and new issues, making calculations of inter-rater reliability meaningless in this exploratory work. We were thus unable to calculate inter-rater reliability for ORBIT assessments as we had planned.

Lastly and most importantly, we did not examine studies excluded from CERs because they did not report data on our outcome of interest ("index outcome"). Such studies are excluded by authors of CERs prior to the synthesis phase of the systematic review, and we did not seek out such studies. We also did not look at other trials in each CER to see if they could have reported out index outcome (but did not). This latter approach was taken by Kirkham and colleagues,[3] enabling them to identify ORBIT classes G and H. Such studies may have exhibited SOR leading to exclusion of the study from a CER, and thus our estimates would underestimate the frequency of SOR. Examination of all studies in each CER and of additional studies reporting on each intervention as identified through registries or FDA documents might lead to a more complete and accurate assessment of SOR for each index outcome. In addition, the conclusions in a CER are based on an assessment of multiple outcomes, both benefits and harms, and we only examined single outcome for each CER.

And finally, our study does not examine the broader issue of bias in study design and choice of outcomes based on anticipated findings by the trialists. This manipulation of study design is not technically SOR or SAR, which are defined in terms of the selective reporting of outcomes based on results. Efforts such as the Core Outcome Measures in Effectiveness Trials (COMET) Initiative[46] are working to address this issue. Although such design issues are a critical source of bias affecting the internal validity of trials, they are beyond the scope of this report.

Challenges Encountered Developing Our Study Protocol

The methodology of this project changed significantly as it proceeded. We had initially proposed exploring SOR and SAR in nonrandomized studies, to develop an ORBIT-like classification system for nonrandomized studies, and to quantify the effects of SOR/SAR on effect estimates in meta-analyses in CERs. These initial goals proved infeasible, however, and were changed after lengthy discussions among the coauthors of this report. In the absence of registry information or any other prespecified characteristics on nonrandomized studies, we did not feel that we were able to explore the frequency of SOR/SAR in these study designs. We felt that determination of SOR/SAR based purely on information in the publication would not be useful.

We therefore evolved our objectives to focus on RCTs, with an assessment of the proportion of trials included in CERs with SOR or SAR, determination of the inter-rater reliability of the ORBIT classification system, and an examination of ORBIT's validity when used to assess trial publications (as the tool was intended) when compared with ORBIT assessment using trial registry information. ORBIT proved inadequate for our assessment of the types of SOR and SAR that we identified in included RCTs, and assessment of ORBIT's inter-rater reliability and validity proved infeasible as discussed above.

We also explored quantifying the effect of SOR/SAR in RCTs by comparing the effect estimate from meta-analyses in CERs to estimates obtained using imputed data in studies with missing outcome data (i.e., SOR), using an approach such as that of Copas and Jackson.[8] After

discussion among the coauthors of this report, we decided that such analyses would not be useful in view of the multitude of assumptions made in such imputations, and the difficulty determining if SOR/SAR existed in individual studies. In addition, studies contributing to pooled estimates frequently included those published both before and after 2005. The former studies were rarely listed in trial registries and more recent studies were not always registered. The coauthors of this report felt that we should focus on describing SOR/SAR and exploring ORBIT, rather than on problematic quantitative estimates.

We had initially proposed examining SOR and SAR separately. The distinction between the two was often unclear, however. For example, an outcome such as A1c could be analyzed in several ways, including as a continuous measure (with absolute change in percent) or as categorical outcome with various thresholds defining the categories. These presentations of A1c could be considered different outcomes or different analyses. Thus the distinction between SOR and SAR is somewhat arbitrary.

Challenges and Recommendations Regarding Methods and Available Tools for Exploring Selective Outcome Reporting and Selective Analysis Reporting

As a result of our exploration of SOR and SAR in a small cohort of RCTs and our use of trial registries to identify SOR/SAR and to assign an ORBIT class, we formulated a number of comments and suggestions on these tools and the potential approaches that systematic reviewers might use (Box 1). Our suggestions are intended to stimulate and provide a basis for future discussions that might ultimately lead to the development of explicit guidance for systematic reviewers.

ORBIT

Challenges and Limitations of the ORBIT Classification

We encountered a number of problems when we used ORBIT to assess included trials for SOR and SAR. First, we had difficulty making distinctions among ORBIT categories. Nine categories are a large number to have assessors consider with an adequate degree of reliability. Our team's assessors had difficulty making the distinction between categories E and F, for example, which depends upon an assessment of why an outcome might have been measured but not reported. Second, in some studies it was difficult to determine if an exact p value could be determined from the data presented in the trial publication, particularly if the assessor did not have a statistical background. Kirkham and colleagues[3] indicate that determination of the ORBIT class can use data calculated indirectly from the results (e.g., an exact p value calculated from the standard error of an estimate). With adjusted and between-group analyses it was sometimes unclear to the assessor whether an exact p value could be calculated without access to the underlying dataset.

The most important issues that we encountered with ORBIT were not with its implementation, but rather with the limited nature of its intended use and scope. ORBIT was designed for the assessment of SOR/SAR using information within the publication(s) plus clinical judgment.[3] We assert, however, that the most important indicators of SOR and SAR are obtained from sources outside of the published report, such as trial registries or databases of research protocols. ORBIT does not incorporate this additional information. Our efforts to

compare information in the trial registry, including results (if any), with the publication (judgments #2 and #3) were met with limited success because the most common situation was the addition of outcomes that were not prespecified, and ORBIT does not accommodate that type of SOR or SAR. Registry information could contribute to assessments of ORBIT classes G and H – when it was unclear in the publication if the outcome was measured, but we did not encounter G and H assessments because of our study design.

In addition, ORBIT only addresses missing or incomplete outcome reporting as it considers outcomes that the reader is led to believe will be in the results section - either because they are mentioned in the methods section, or clinical judgment or information from other studies suggests that an outcome should be reported. The far more frequent scenario in our experience, however, was the reporting of outcomes in publications that had not been mentioned either in the methods section of the publication or in the trial registry. ORBIT also does not include the frequent changes that appear to occur in the primary or secondary outcomes in a study: either when or how an outcome was measured, or the evolution of an outcome initially poorly specified (e.g., "glycemic control") to one with specificity after subject recruitment had been completed. Such changes can only be identified by a careful review of the History of Changes in ClinicalTrials.gov. ORBIT also does not address the issue of the validity of study conclusions based on data reported in a given publication. Studies have been documented to present biased conclusions that correlate with the funders' interests.[37, 39, 40, 47]

Perhaps most importantly, because ORBIT focuses exclusively on the study publication(s), poor writing can lead directly to the appearance of SOR and SAR, where the "selectively" reported outcome is missing simply because of an error or omission or because of constraints on the number of words in the publication. In this situation, apparent SOR and SAR might not lead to actual bias in reporting, if the selective reporting does not relate to the direction and statistical significance of the results. On the other hand, SOR and SAR (and the presence of outcome reporting bias) could be obscured by careful writing.

These limitations in the scope of ORBIT as a tool for assessing SOR/SAR are not criticisms of ORBIT per se, as it was not intended to address these additional types of selective reporting of outcomes and analyses. These limitations, however, point to the need for additional tools to identify and assess SOR/SAR in primary studies.

Recommendations for Future Research on the ORBIT Classification System

A new tool is needed for the assessment of SOR/SAR based on all available information, both within and beyond the study publication(s), including trial registries and protocols and unpublished data sources. Such a tool needs to have a limited number of categories and should be broadly applicable to a variety of study interventions and outcomes. This new instrument needs to incorporate both the selection of outcomes from those prespecified, and the addition of new outcomes that differ from those prespecified in the methods section or trial registry. The types of SOR/SAR incorporated in our judgments #2 and #3 (Table 3) may be a useful starting point for development of such a tool. Based on our experiences with the ORBIT classification system, we suggest that the number of categories be relatively small, and judgments based on undocumented information be removed (e.g., the distinction between ORBIT classes E and F, and G and H). Such a tool needs early evaluation, including both reliability and validity assessments.

At a later date, classification systems for nonrandomized studies need to be developed also. We recommend, however, that systems for RCTs be the initial focus for research, as trials are more uniform in design, more information is available in registries and other databases, and nonrandomized studies are often exploratory – purposefully without prespecification of all outcomes and analyses.

If ORBIT continues to be used, research is needed to determine the intra- and inter-rater reliability of the ORBIT classification system. To our knowledge, there are no reliability data on this instrument. A cohort of individuals with training and experience in the critical appraisal of trials should examine studies from a variety of subject fields and with a variety of index outcomes. Efficacy, effectiveness, and harms outcomes should be examined for SOR/SAR, as the frequency of SOR and SAR and the reliability of the ORBIT tool likely vary across types of outcomes.

In addition, the ORBIT classification system needs to be validated. Kirkham and coauthors[3] performed a limited assessment of the accuracy of ORBIT classes G and H (unclear whether the outcome was measured or not) by comparing their designated class to information provided by the trialists. The sensitivity for predicting that the outcome had been measured was 92 percent and the specificity for predicting that the outcome had not been measured was 77 percent. This assessment of the accuracy of ORBIT was calculated based on a response rate from the trialists of 12 percent (65 of 538 author reports).

Trial Registries

Problems Encountered Using Trial Registries

Trial registries are an important recent advancement in biomedical research, improving public knowledge about ongoing and completed trials, promoting access to research results, and delineating prespecified study methods. Study registries were not as useful as we had anticipated for identifying and characterizing SOR and SAR, however. The following comments focus on ClinicalTrials.gov, as that was the predominant registry that we encountered, other registries that we examined contained less information than did ClinicalTrials.gov, and only this registry contains results (Box 1).

The list of primary and secondary study outcomes on the summary page for each study in ClinicalTrials.gov often did not meet our needs or it was unclear how best to use the information that was provided. ClinicalTrials.gov indicates the "original" and "current" outcomes, both primary and secondary. This information can be used to quickly determine if there was a change in the primary and secondary outcomes and/or a change in an outcome's designation as primary or secondary. This information was most useful when outcomes were fully specified in terms of how and when they were measured. Frequently, however, outcomes listed in the registry were inadequately specified, particularly the original outcome, thus we could not determine if there was a significant change between the original prespecification and the published outcomes. For example, "A1c" might be specified as the original outcome in the registry, and the current outcome listed as "change in A1c (percent) from baseline to 26 weeks." Although we classified this as a change in the index outcome and considered it potential SOR, the trialists' original outcomes and their motives for the change were unclear.

In addition, the summary page with the "original" and "current" outcomes was only part of the story on changes in specified outcomes over the course of a trial. Additional information could be found by clicking on the "History of Changes" link, where various aspects of study

outcomes could change numerous times during the course of the trial, but only the "original" and "current" outcomes are captured on the registry summary page. It is very time consuming to review the "History of Changes," and difficult to determine if changes represented potential SOR. Although one might suspect SOR/SAR based on the timing of changes, it was not possible to determine if the trialists were purposefully manipulating reported outcomes after they had performed analyses and identified outcomes that they considered favorable.

Safety outcomes appeared to be less precisely specified in the registry, although our experiences were confined to the outcome of mortality for the CER on lipid agents.[24] In reviewing the other two CERs, we rarely encountered any specification of safety outcomes, although we did not quantify those findings.

The study information sections of trial registries provided no information informing SAR for any of our included studies. No trial registration provided any information on analyses set, subgroups, proposed analyses, or covariates. The registry results frequently provided such information, but since that information was entered after analyses were completed, it was not prespecified and thus did not inform our assessment of SAR.

Registry results, when available, did provide useful information that could be compared to the trial publication. In the majority of RCTs with available results, however, the registry results were less complete than the publication for our index outcome. In other words, additional analyses and results were published that were not presented in the registry, while we rarely encountered the situation where the registry results were more complete. We consider the addition of new outcomes and analyses an important type of SOR/SAR, because they may reflect results favorable to the authors and/or study funders.

Frequently results were not presented in the registry (53 percent of our studies overall). There are legitimate reasons why a trial might not have results posted in a registry.[48] ClinicalTrials.gov, however, does not provide a reason when trial results are not posted in the registry, and only the situation of an ongoing study can be deduced from the registry.

For each included RCT registered with ClinicalTrials.gov, we examined the "History of Changes" page of the registry for relevant changes and encountered a number of challenges. First, the vast majority of changes were of no relevance to our exploration of SOR and SAR, and it was difficult to efficiently identify potentially relevant changes among the plethora of information. The "History of Changes" contains spelling and punctuation corrections, the addition of abbreviations, the addition of citations for publication derived from the trial, the addition of study results, and changes in outcomes, among many other types of changes. Although there are categories of changes in ClinicalTrials.gov, including protocol, recruitment status, location/contact, administrative, and miscellaneous, protocol changes were often combined with administrative changes if they occurred on the same date. Thus the user often had to review irrelevant information when looking for significant protocol changes.

Second, in addition to the amount of information that had to be reviewed, the format for presentation of changes was suboptimal for our purpose. For example, dates for each change in outcome were specified, however, how that date related to important events such as initial or final subject recruitment was not transparent (although all relevant dates could be determined and compared with significant effort). In addition, all changes were formatted as fields for data entry (e.g., "<textblock>") and using abbreviations that were not defined. Such a format is not user-friendly and the meaning was unclear at times. When results were added to the registry, they were listed also on the "History of Changes" page, but without any formatting and with

their data entry code: a format of little use to a reader. (The results were clearly presented in the "Study Results" tab, however).

Preliminary Recommendations for Changes in ClinicalTrials.gov

From our experience trying to identify and characterize SOR and SAR from information in trial registries, and from using ClinicalTrials.gov in particular, we present a number of suggestions for improving ClinicalTrials.gov and registries in general (Box 2). These suggestions are based solely on this exploratory work, and they need to be further vetted and validated before being implemented. Our suggestions are of two types: the first related to presentation and ease of use of the information contained in the registry; and second to the content of the registry. We will address both types, focusing only on ClinicalTrials.gov in view of its dominance in our study. Our suggestions are confined to the use of a trial registry for obtaining information on potential SOR and SAR, and are not intended to encompass other reasons for using a trial registry.

Changes in Format of ClinicalTrials.gov

A visual timeline of relevant dates would be useful to users of ClinicalTrials.gov. Currently, numerous dates are presented in narrative form in the introductory page and/or the "History of Changes" page. For many dates there is a logical sequence that should occur if bias in the trial is to be minimized. For example, trial registration should always occur prior to commencement of subject recruitment. The presentation of a timeline of critical points in trial design, implementation, analysis, and publication would allow the user to quickly assess whether the chronology of study design and registration was optimal.

The formatting of the "History of Changes" page needs extensive revision to be of optimal use. The user should be able to review types of changes without having to review categories that are not of interest. For example, if the user is only interested in changes in trial protocol methodology such as a change in the primary outcome or followup interval, the user should not have to review spelling corrections that were entered on the same date as the change in study protocol. The current categories of changes are reasonable; however the approach of organizing by date of change is not. In addition, the field codes should be changed to meaningful labels and headings, or eliminated completely. In addition, all listed changes should represent real changes: situations were encountered where a change was listed, but the text before and after appeared identical.

The "History of Changes" page as currently formatted is not appropriate for reporting results. Results are clearly and efficiently presented in the results section: the "History of Changes" could refer the reader there, with documentation of the date of addition of results to the registry.

Changes in Content of ClinicalTrials.gov

Primary and secondary outcomes need to be specified in detail in the registry: what the outcome is, how it will be measured (if appropriate), and when it will be measured. For example the outcome A1c needs to be fully qualified, such as "between-group change in A1c (percent) from baseline to 26 weeks." Listing of vague outcomes like "glycemic control" should not ever be permitted in trial registries, particularly after subject recruitment has begun. Precise prespecification of outcomes needs to apply to both "original" and "current," primary and secondary, and benefits and harms outcomes. Zarin and colleagues in 2011[34] provide a useful framework for describing the levels of specification of outcomes measures. Ideally trial registries

would require that outcomes be prespecified at "level 4, Method of Aggregation," which refers to how the variable was measured, and the specific measure and timepoint (e.g., change in A1c measured as a continuous outcome at 26 weeks followup and reported as mean change).

Descriptions of changes in outcomes, both primary and secondary, should be readily accessible to the user. Ideally, in addition to the "original" and "current" outcomes now displayed on the trial's main page, a chronology of all changes in outcomes would be clearly presented. The trialist should also be required to indicate the reason for any change in primary or secondary outcomes that were made after subject recruitment commenced. Substantive changes after that point are rarely indicated, and thus when they occur the study authors need to provide a rationale. Population, intervention, and other important subgroups that are part of confirmatory (vs. exploratory) analyses should also be specified in the registry prior to commencing subject recruitment.

For the purposes of identifying SAR, it would be very useful for trial registries to include prespecification of selected information such as the general analytic approach (superiority, noninferiority, equivalence) and the analysis set (e.g., full analysis set and how that was defined). Other information, such as covariates for adjusted analyses would also be useful.

Ideally trial registries would contain the full study protocols, as has been suggested by other researchers.[49] These detailed documents, finalized prior to the start of subject recruitment and only modified with specific and explicit justification, document the study design and analytic approaches that can be used to assess SOR and SAR.

Mechanisms for enforcing registration and the completeness and accuracy of data entries may be required to achieve high quality registries that are most useful to patients, researchers, and systematic reviewers, as has been suggested by others.[50]

Box 1. ORBIT: Challenges and limitations

Challenges using the ORBIT classification system

1. Difficulty making distinctions among the ORBIT categories
2. Problems determining if an exact p value could be determined from the data presented

Limitations of the ORBIT classification system

1. Scope and intended use: ORBIT was designed for use with the trial publication(s) only, and not for use with additional information such as that obtained from trial registries or protocols. We found that this additional information could rarely be applied to ORBIT categories.
2. ORBIT only addresses missing outcomes, i.e., those that the reader expects from reading the methods section or from clinical judgment. ORBIT does not include the addition of outcomes or analyses to results section that do not appear to have been specified a priori.

Box 2. ClinicalTrials.gov: Problems and recommendations for changes

Problems encountered using ClinicalTrials.gov

1. Outcomes listed on the summary page were frequently inadequately specified (e.g., "glycemic control")
2. Only "original" and "current" primary and secondary outcomes are listed, although numerous changes in outcomes can have occurred between those two listings.
3. The "History of Changes" tab was difficult to use: labels were unclear and it was difficult to efficiently identify important changes in the outcomes.
4. Safety outcomes were rarely specified in the summary page, and when they were listed often lack specificity.
5. No information of use to detect SAR is provided, such as the analysis set or variables for adjustment.
6. Results were often not posted in the registry and when they were, they were often less complete than those in the publication.

Recommended Changes in ClinicalTrials.gov

1. Changes in format
    a. A visual timeline of important dates would assist the reader in evaluating when registration and changes in outcomes occurred with respect to subject recruitment and data analysis.
    b. The "History of Changes" pages should be formatted to facilitate efficient identification of important protocol changes and data labels should be self-explanatory.
2. Changes in content
    a. Primary and secondary outcomes should be specified in detail: what, how, and when.
    b. Any changes in outcomes, both primary and secondary, should be listed with the date and rationale on the registry summary page.
    c. Information on planned analytic approach (e.g., superiority, non-inferiority), the analysis set (e.g., intention-to-treat, as-treated), and variables planned for adjusted analyses should be provided.

Suggestions for Identifying and Characterizing Selective Outcome Reporting and Selective Analysis Reporting

As a result of this exploratory work, we developed several suggestions for systematic reviewers to use when trying to determine the frequency and effect of SOR and SAR during a systematic review. These suggestions will inform further discussions as guidance is formulated in future. In view of the paucity of evidence on the prevalence and effect of SOR and SAR on effect estimates and conclusions in systematic reviews, our suggestions should also inform future research agendas.

Most importantly, systematic reviewers need to encompass a broad definition of SOR and SAR, including not only the situation where a subset of the original outcomes measured and analyzed in a trial are fully reported based on the magnitude of the treatment effect or the statistical significance of selected outcomes, but also [2]the publication of outcomes and analyses that were not prespecified in the publication methods section or in the trial's registration information. The reporting of this expanded set of outcomes also represents the selective reporting of outcomes or analyses, likely based on the direction, magnitude, or statistical significance of the results.

Implications for Systematic Reviewers

The systematic review team needs to be strategic and parsimonious in their efforts to identify SOR and SAR. We suggest that the team consider the following steps in their assessment of potential SOR and SAR:

1. Review trial publication(s) for incompletely reported outcomes in the results section (i.e., ORBIT classes A through D). Although these classes may be infrequent and may also represent writing and journal styles and constraints on the number of words in a publication, they are important classes to identify. We do not suggest trying to identify and categorize the other ORBIT classes by examining only the study publication because the discrepancies identified may reflect for the most part poor writing and not necessarily the selective reporting of outcomes based on the nature and direction of the results.
2. The systematic reviewer should not routinely seek information from trial registries for the purpose of identifying SOR or SAR in every study included in a review. This step is potentially very resource intensive and the gain in understanding of SOR and SAR is likely to be minimal. In addition, we were not able to identify predictors of SOR/SAR in our small cohort of studies that had little variation in several of the potential predictors that we examined. Until further research is available to inform such predictors, we offer the following suggestions based on our experience.
    a. If there is a concern about SOR in a trial publication, either because of incomplete reporting of outcomes, missing outcomes that likely should have been reported, or other clues, trial registration should be sought out, and if identified, the registry should be searched for reference to a trial protocol and for results posted to the registry.
    b. If the trial registry does not include results or reference to a study protocol, we do not recommend further exploration of the summary information provided in the registry because the additional information is unlikely to be helpful.

    c. In the absence of specific reference to study registration in the publication, only search registries for RCTs. Nonrandomized studies are not registered frequently enough for searches to be worthwhile.
    d. Given the current focus of registries on outcomes of benefit, at present we suggest searching only for outcomes of benefit, and not harms outcomes unless the latter were likely to have been prespecified as the trial's primary or secondary outcome. In our experience, useful prespecification of safety outcomes was even less common than for efficacy and effectiveness outcomes.
3. When searching for registered trials, use the World Health Organization ICTRP portal. This search engine accesses a number of different trial registries at once including ClinicalTrials.gov.
4. Comparison of the methods and results sections of trial publications will rarely provide useful information for identifying and characterizing SOR or SAR. The methods section of publications is most likely written after data analyses were completed, and discrepancies between the methods and results sections likely represent poor writing on the part of the trial's authors and discrepancies cannot be attributed with certitude to SOR or SAR.

Conclusions

In this exploratory study of a small cohort of RCTs that were included in three CERs, we determined that trials published in or after 2005 and contributing to meta-analyses, were not consistently listed in trial registries. Publications of trials do not consistently report information on trial registration, and reporting of results in ClinicalTrials.gov was inconsistent with no clear reasons for the inconsistencies.

We identified numerous challenges in searching for, and characterizing, SOR and SAR in our cohort of RCTs. We did not find the ORBIT classification tool,[3] designed for the assessment of SOR within trial publications, particularly useful. ORBIT has too many categories with too much ambiguity among those categories. ORBIT classes did not describe the types of SOR and SAR that we most frequently encountered: the addition of outcomes measures, subgroups, and other analyses to published results that were not prespecified in the publication methods section or listed in the registry. We consider this type of SOR and SAR as important as the nonreporting of prespecified outcomes.

Trial registries were of little use in identifying SOR unless trial results were listed in the registry, given the current lack of specificity of outcome designation in registries. Registries were of no use in identifying SAR. The presentation and content of ClinicalTrials.gov, the predominant registry for the trials that we examined, could be improved to better assist the systematic reviewer in identifying potential SOR and SAR. Suggestions for improvements in trial registries, and ClinicalTrials.gov in particular, include: (1) a requirement that outcomes be precisely specified; (2) the reasons for any change in primary or secondary outcomes should be provided by the trialist; (3) improved formatting of the History of Changes section; (4) prespecification of the analysis set and general analytic approach be mandated; and (5) efforts to insure that all available results are posted.

Much further research is needed to develop efficient, tailored approaches to identifying and characterizing SOR and SAR in trials starting with an expanded and simplified classification system and changes to trials registries. Without such improvements, the increased time needed by systematic reviewers to try and identify and characterize SOR and SAR are not likely to be worthwhile. Research ultimately needs to guide systematic reviewers in assessing the direction and magnitude of the effects of missing outcomes and analyses on effect estimates and conclusions in systematic reviews.

References

1. Sterne J, Egger M, Moher D. Addressing reporting biases. Cochrane Handbook for Systematic Reviews of Interventions. 2008;Chichester: John Wiley & Sons Ltd:297-333.

2. Hutton JL, Williamson PR. Bias in meta-analysis due to outcome variable selection within studies. Applied Statistics 2000. 2002;49(3):359-70.

3. Kirkham JJ, Dwan KM, Altman DG, et al. The impact of outcome reporting bias in randomised controlled trials on a cohort of systematic reviews. BMJ. 2010;340:c365. PMID: 20156912.

4. Higgins J, Altman D. Assessing risk of bias in included studies. Cochrane Handbook for Systematic Reviews of Interventions. 2008;Chichester: John Wiley & Sons Ltd:187-242.

5. Dickersin K, Chalmers I. Recognising, investigating and dealing with incomplete and biased reporting of clinical research: from Francis Bacon to the World Health Organisation. James Lind Library. 2010.

6. Dwan K, Altman DG, Arnaiz JA, et al. Systematic Review of the Empirical Evidence of Study Publication Bias and Outcome Reporting Bias. PLoS ONE. 2008;3(8):e3081.

7. Song F, Parekh S, Hooper L, et al. Dissemination and publication of research findings: an updated review of related biases. Health Technol Assess 2010;14(8):1-220.

8. Copas J, Jackson D. A Bound for Publication Bias Based on the Fraction of Unpublished Studies. Biometrics. 2004;60(1):146-53.

9. Hart B, Lundh A, Bero L, et al. Effect of reporting bias on meta-analyses of drug trials: reanalysis of meta-analyses. BMJ. 2012;344:d7202. PMID: 22214754.

10. Dickersin K. The existence of publication bias and risk factors for its occurrence. JAMA. 1990 Mar 9;263(10):1385-9. PMID: 2406472.

11. Dickersin K, Chan S, Chalmers TC, et al. Publication bias and clinical trials. Controlled Clinical Trials. 1987 Dec;8(4):343-53. PMID: 3442991.

12. Hopewell S, Loudon K, Clarke M. Publication bias in clinical trials due to statistical significance or direction of trial results. Cochrane Database Syst Rev. 2009;1:MR000006.

13. Stern JM, Simes RJ. Publication bias: evidence of delayed publication in a cohort study of clinical research projects. BMJ. 1997 Sep 13;315(7109):640-5. PMID: 9310565.

14. Dickersin K, Olson CM, Rennie D, et al. Association between time interval to publication and statistical significance. JAMA. 2002 Jun 5;287(21):2829-31. PMID: 12038925.

15. Gotzsche PC. Reference bias in reports of drug trials. British Medical Journal Clinical Research Ed. 1987 Sep 12;295(6599):654-6. PMID: 3117277.

16. Food and Drug Administration Modernization Act of 1997. 1997;Pub L No. 105-115.

17. United States Code. US Public Law 110–85: Food and Drug Administration Amendments Act of 2007. 2008. http://frwebgate.access.gpo.gov/cgi-bin/getdoc.cgi?dbname=110_cong_public_laws&docid=f:publ085.110. Accessed on May 7, 2012.

18. World Health Organization. International Clinical Trials Registry Platform (ICTRP). 2011. www.who.int/ictrp/en/. Accessed on August 9, 2011.

19. Bennett W, Wilson L, Bolen S, et al. Oral Diabetes Medications for Adults With Type 2 Diabetes: An Update. Comparative Effectiveness Review No. 27. (Prepared by Johns Hopkins University Evidence-based Practice Center under Contract No. 290-02-0018.) AHRQ Publication No. 11-EHC038-EF. Rockville, MD: Agency for Healthcare Research and Quality. March 2011. www.effectivehealthcare.ahrq.gov/reports/final.cfm

20. Phung OJ, Coleman CI, Baker EL, et al. Effectiveness of Recombinant Human Growth Hormone (rhGH) in the Treatment of Patients With Cystic Fibrosis. Comparative Effectiveness Review No. 23. (Prepared by the University of Connecticut/Hartford Evidence-based Practice Center under Contract No. 290-2007-10067-I) AHRQ Publication No. 11-EHC003. Rockville, MD: Agency for Healthcare Research and Quality. October 2010. www.effectivehealthcare.ahrq.gov/reports/final.cfm.

21. Yank V, Tuohy CV, Logan AC, et al. Comparative Effectiveness of Recombinant Factor VIIa for Off-Label Indications vs. Usual Care. Comparative Effectiveness Review No. 21. (Prepared by Stanford-UCSF Evidence-based Practice Center under Contract No. #290-02-0017) Rockville, MD: Agency for Healthcare Research and Quality. May 2010. www.effectivehealthcare.ahrq.gov/reports/final.cfm.

22. Coleman CI, Baker WL, Kluger J, et al. Comparative Effectiveness of Angiotensin Converting Enzyme Inhibitors or Angiotensin II Receptor Blockers Added to Standard Medical Therapy for Treating Stable Ischemic Heart Disease. (Prepared by the University of Connecticut/Hartford Hospital Evidence-based Practice Center under Contract No. 290-2007-10067-I.) Rockville, MD: Agency for Healthcare Research and Quality. October 2009. www.effectivehealthcare.ahrq.gov/reports/final.cfm.

23. Nelson HD, Fu R, Humphrey L, et al. Comparative Effectiveness of Medications To Reduce Risk of Primary Breast Cancer in Women. Comparative Effectiveness Review No. 17. (Prepared by Oregon Evidence-based Practice Center under Contract No. 290-2007-10057-1.) Rockville, MD: Agency for Healthcare Research and Quality. September 2009. www.effectivehealthcare.ahrq.gov/reports/final.cfm.

24. Sharma M, Ansari M, Soares-Weiser K, et al. Comparative Effectiveness of Lipid-Modifying Agents. Comparative Effectiveness Review No. 16. (Prepared by the University of Ottawa Evidence-based Practice Center under contract No. 290-02-0021.) Rockville, MD: Agency for Healthcare Research and Quality. September 2009. www.effectivehealthcare.ahrq.gov/reports/final.cfm.

25. Qayyum R, Wilson LM, Bolen S, et al. Comparative Effectiveness, Safety, and Indications of Insulin Analogues in Premixed Formulations for Adults with Type 2 Diabetes. Comparative Effectiveness Review No. 14. (Prepared by the Johns Hopkins University Evidence-based Practice Center under Contract No. 290-02-0018.) Rockville, MD: Agency for Healthcare Research and Quality. September 2008. www.effectivehealthcare.ahrq.gov/reports/final.cfm.

26. MacLean C, Alexander A, Carter J, et al. Comparative Effectiveness of Treatments To Prevent Fractures in Men and Women With Low Bone Density or Osteoporosis. Comparative Effectiveness Review No. 12. (Prepared by Southern California/RAND Evidence-based Practice Center under Contract No. 290-02-0003). Rockville, MD: Agency for Healthcare Research and Quality. December 2007. www.effectivehealthcare.ahrq.gov/reports/final.cfm.

27. Donahue KE, Gartlehner G, Jonas DE, et al. Comparative Effectiveness of Drug Therapy for Rheumatoid Arthritis and Psoriatic Arthritis in Adults. Comparative Effectiveness Review No. 11. (Prepared by RTI-University of North Carolina Evidence-based Practice Center under Contract No. 290-02-0016.) Rockville, MD: Agency for Healthcare Research and Quality. November 2007. www.effectivehealthcare.ahrq.gov/reports/final.cfm.

28. Matchar DB, McCrory DC, Orlando LA, et al. Comparative Effectiveness of Angiotensin-Converting Enzyme Inhibitors (ACEIs) and Angiotensin II Receptor Antagonists (ARBs) for Treating Essential Hypertension. Comparative Effectiveness Review No. 10. (Prepared by Duke Evidence-based Practice Center under Contract No. 290-02-0025.) Rockville, MD: Agency for Healthcare Research and Quality. November 2007. www.effectivehealthcare.ahrq.gov/reports/final.cfm.

29. Gartlehner G, Hansen RA, Thieda P, et al. Comparative Effectiveness of Second-Generation Antidepressants in the Pharmacologic Treatment of Adult Depression. Comparative Effectiveness Review No. 7. (Prepared by RTI International-University of North Carolina Evidence-based Practice Center under Contract No. 290-02-0016.) Rockville, MD: Agency for Healthcare Research and Quality. January 2007. www.effectivehealthcare.ahrq.gov/reports/final.cfm.

30. Shekelle P, Maglione M, Bagley S, et al. Comparative Effectiveness of Off-Label Use of Atypical Antipsychotics. Comparative Effectiveness Review No. 6. (Prepared by the Southern California/RAND Evidence-based Practice Center under Contract No. 290-02-0003.) Rockville, MD: Agency for Healthcare Research and Quality. January 2007. www.effectivehealthcare.ahrq.gov/reports/final.cfm.

31. Chou R, Helfand M, Peterson K, et al. Comparative Effectiveness and Safety of Analgesics for Osteoarthritis. Comparative Effectiveness Review No. 4. (Prepared by the Oregon Evidence-based Practice Center under Contract No. 290-02-0024.) Rockville, MD: Agency for Healthcare Research and Quality. September 2006. www.effectivehealthcare.ahrq.gov/reports/final.cfm.

32. Seidenfeld J, Piper M, Bohlius J, et al. Comparative Effectiveness of Epoetin and Darbepoetin for Managing Anemia in Patients Undergoing Cancer Treatment. Comparative Effectiveness Review No. 3. (Prepared by Blue Cross and Blue Shield Association Technology Evaluation Center Evidence-based Practice Center under Contract No. 290-02- 0026.) Rockville, MD: Agency for Healthcare Research and Quality. May 2006. www.effectivehealthcare.ahrq.gov/reports/final.cfm

33. Ip S, Bonis P, Tatsioni A, et al. Comparative Effectiveness of Management Strategies for Gastroesophageal Reflux Disease. Comparative Effectiveness Review No. 1. (Prepared by Tufts-New England Medical Center Evidence-based Practice Center under Contract No. 290-02-0022.) Rockville, MD: Agency for Healthcare Research and Quality. December 2005. www.effectivehealthcare.ahrq.gov/reports/final.cfm.

34. Zarin DA, Tse T, Williams RJ, et al. The ClinicalTrials.gov results database--update and key issues. New England Journal of Medicine. 2011 Mar 3;364(9):852-60. PMID: 21366476.

35. FDA Failed to Enforce Law Requiring Drugmakers to Disclose Test Data- WashPost. 2004.

36. Zarin DA, Ide NC, Tse T, et al. Issues in the registration of clinical trials. JAMA. 2007 May 16;297(19):2112-20. PMID: 17507347.

37. Als-Nielsen B, Chen W, Gluud C, et al. Association of funding and conclusions in randomized drug trials: a reflection of treatment effect or adverse events? JAMA. 2003 Aug 20;290(7):921-8. PMID: 12928469.

38. Bekelman JE, Li Y, Gross CP. Scope and Impact of Financial Conflicts of Interest in Biomedical Research. JAMA: The Journal of the American Medical Association. 2003 January 22, 2003;289(4):454-65.

39. Lexchin J, Bero LA, Djulbegovic B, et al. Pharmaceutical industry sponsorship and research outcome and quality: systematic review. BMJ. 2003 May 31;326(7400):1167-70. PMID: 12775614.

40. Sismondo S, Sismondo S. How pharmaceutical industry funding affects trial outcomes: causal structures and responses. Social Science & Medicine. 2008 May;66(9):1909-14. PMID: 18299169.

41. Raz I, Chen Y, Wu M, et al. Efficacy and safety of sitagliptin added to ongoing metformin therapy in patients with type 2 diabetes. Current Medical Research & Opinion. 2008 Feb;24(2):537-50. PMID: 18194595.

42. Scott R, Loeys T, Davies MJ, et al. Efficacy and safety of sitagliptin when added to ongoing metformin therapy in patients with type 2 diabetes. Diabetes, Obesity & Metabolism. 2008 Sep;10(10):959-69. PMID: 18201203.

43. Rosenstock J, Rood J, Cobitz A, et al. Initial treatment with rosiglitazone/metformin fixed-dose combination therapy compared with monotherapy with either rosiglitazone or metformin in patients with uncontrolled type 2 diabetes. Diabetes, Obesity & Metabolism. 2006 Nov;8(6):650-60. PMID: 17026489.

44. Easterbrook PJ, Berlin JA, Gopalan R, et al. Publication bias in clinical research. Lancet. 1991 Apr 13;337(8746):867-72. PMID: 1672966.

45. Bekkering GE, Harris RJ, Thomas S, et al. How much of the data published in observational studies of the association between diet and prostate or bladder cancer is usable for meta-analysis? American Journal of Epidemiology. 2008 May 1;167(9):1017-26. PMID: 18403406.

46. Williamson P, Clarke M. The COMET (Core Outcome Measures in Effectiveness Trials) Initiative: its role in improving Cochrane Reviews [editorial]. The Cochrane Library, Accessed May 7, 2012; 2012. www.thecochranelibrary.com/details/editorial/1797057/The-COMET-Core-Outcome-Measures-in-Effectiveness-Trials-Initiative-its-role-in-i.html. Accessed on May 7, 2012.

47. Golder S, Loke YK. Is there evidence for biased reporting of published adverse effects data in pharmaceutical industry-funded studies? British Journal of Clinical Pharmacology. 2008 Dec;66(6):767-73. PMID: 18754841.

48. About the ClinicalTrials.gov Results Database. 2011. http://clinicaltrials.gov/ct2/info/results. Accessed on January 5, 2012.

49. Chan AW, Chan A-W. Out of sight but not out of mind: how to search for unpublished clinical trial evidence. BMJ. 2012;344:d8013. PMID: 22214892.

50. Mathieu S, Boutron I, Moher D, et al. Comparison of registered and published primary outcomes in randomized controlled trials.[Erratum appears in JAMA. 2009 Oct 14;302(14):1532]. JAMA. 2009 Sep 2;302(9):977-84. PMID: 19724045.

Abbreviations and Acronyms

A1c	hemoglobin A1c
ARB	analysis reporting bias
AHRQ	Agency for Healthcare Research and Quality
CERs	comparative effectiveness reviews
EPC	Evidence-based Practice Centers
FDA	United States Food and Drug Administration
ICMJE	International Committee of Medical Journal Editors
ICTRP	International Clinical Trials Registry Platform
IQR	interquartile range
NA	not applicable
NR	not reported
OR	odds ratio
ORB	outcome reporting bias
ORBIT	Outcome Reporting Bias in Trials
RCT	randomized controlled trial
SAR	selective outcome reporting
SOR	selective analysis reporting
WHO	World Health Organization

Appendix A. Trial Registries

The registration of studies, particularly randomized controlled trials (RCTs), is an important tool for identifying all studies related to key question in a comparative effectiveness review (CER). Registries are also a potential tool for assessing SOR and SAR. In the United States, the U.S. Food and Drug Administration (FDA) Modernization Act of 1997 called for the creation of ClinicalTrials.gov and mandated registration of all efficacy drug trials for serious or life-threatening diseases and conditions conducted under FDA Investigational New Drug Application regulations.[1] Each record in ClinicalTrials.gov includes summary information on the study protocol, patient recruitment status, and the location of the study site.

The World Health Organization (WHO) initiated a policy in 2006 requiring trial registration of all medical studies that test treatments on patients or healthy volunteers.[2] WHO developed the International Clinical Trials Registry Platform (ICTRP), a global initiative that aims to make information about all clinical trials involving humans publicly available (http://www.who.int/ictrp/network/primary/en/index.html).[2] The ICTRP operates a Search Portal, which provides access to information about ongoing and completed clinical trials from a number of different trial registries. ICTRP is not a trial registry, but rather provides a single platform for access to trial registration data sets provided by a number of different trial registries, including the following:

- Australian New Zealand Clinical Trials Registry
- Brazilian Clinical Trials Registry (ReBec)
- Chinese Clinical Trial Registry
- Clinical Research Information Service - Republic of Korea
- Clinical Trials Registry - India
- ClinicalTrials.gov
- Cuban Public Registry of Clinical Trials
- German Clinical Trials Register
- International Standard Randomized Controlled Trial Number (ISRCTN) Registry
- Iranian Registry of Clinical Trials
- Japan Primary Registries Network
- Pan African Clinical Trial Registry
- Sri Lanka Clinical Trials Registry
- The Netherlands National Trial Register

For the purposes of registration, the ICTRP defines a clinical trial as any research study that prospectively assigns human participants or groups of humans to one or more health-related interventions in order to evaluate the effects on health outcomes.[2]

The International Committee of Medical Journal Editors (ICMJE) established a policy effective July, 2005, that requires prospective trial registration as a condition of publication and delineates the criteria for an acceptable registry.[3] The ICMJE requirement led to a marked increase in registration in ClinicalTrials.gov in 2005.[4] ClinicalTrials.gov is the largest registry accepted by the ICMJE,[4] however, the ICMJE also accepts registration in any of the primary registries that participate in the WHO platform. ICMJE journals accept "retrospective

registration" (registration occurring after subject enrollment started) of trials that began before July 1, 2005. After that date, however, ICMJE considers publication of trials only if registration occurred before the first patient was enrolled ("prospective registration").[5]

Trial registration has further evolved to include the results for some RCTs. The FDA Amendments Act of 2007, effective September, 2008, requires that clinical trial results be made publicly available on the internet in a database of both registry and results, although some exceptions are permitted.[6]

Appendix A References

1. Food and Drug Administration Modernization Act of 1997. 1997;Pub L No. 105-115.

2. World Health Organization. International Clinical Trials Registry Platform (ICTRP). 2011. www.who.int/ictrp/en/. Accessed on August 9, 2011.

3. DeAngelis CD, Drazen JM, Frizelle FA, et al. Clinical trial registration: a statement from the International Committee of Medical Journal Editors. JAMA. 2004 Sep 15;292(11):1363-4. PMID: 15355936.

4. Zarin DA, Ide NC, Tse T, et al. Issues in the registration of clinical trials. JAMA. 2007 May 16;297(19):2112-20. PMID: 17507347.

5. International Committee of Medical Journal Editors. 2009. www.icmje.org/faq_clinical.html. Accessed on August, 9, 2011.

6. United States Code. US Public Law 110–85: Food and Drug Administration Amendments Act of 2007. 2008. http://frwebgate.access.gpo.gov/cgi-bin/getdoc.cgi?dbname=110_cong_public_laws&docid=f:publ085.110. Accessed on May 7, 2012.

Appendix B. Tables and Figures

Table B1. Oral Hypoglycemic Agents: Summary Table

	Outcomes					Publication methods and results								
Study (name, year)	Change in A1c from baseline	% of patients with A1c <7.0%	% of patients with A1c <6.5%	% patients with decrease in A1c > specific value (specify)	Other A1c outcome	Change in primary and/or secondary outcome with respect to A1c between methods and results	Outcome in methods NR in results	Outcome in methods in-adequately reported in results	Outcome in results NR in methods	Change in analyses between methods and results	Subgroups reported in results that were not described in the methods section	Is SOR/ SAR present based on the publica-tion	ORBIT based on publication (judgment 1)	Change in A1c outcome between original and current outcomes listed in registry
Aschner, 2010	Y	Y	Y	N	N	N	N	N	N	N	N	N	NA, no SOR/SAR	N
Bakris, 2006	Y	N	N	N	N	N	N	N	N	N	N	N	NA, no SOR/SAR	N
Bunck, 2009	Y	N	N	N	Y <7.1%	N	N	N	Y	N	N	Y	SOR/SAR but ORBIT does not apply	Y
Defronzo, 2009	Y	Y	N	N	N	N	N	N	N	N	Y	Y	B, D, NA	N
Defronzo, 2010	Y	N	N	N	N	N	N	Y	Y	N	N	Y	B; SOR/SAR but ORBIT does not apply	Y
Garber, 2009	Y	Y	Y	N	N	N	N	N	N	Y	Y	Y	D; SOR/SAR but ORBIT does not apply	N
Goldberg, 2005	Y	N	N	N	N	N	N	N	N	N	N	N	NA, no SOR/SAR	N
Goldstein, 2007 Williams-Herman, 2009	Y	Y	Y	N	N	N	N	N	N	N	N	N	NA, no SOR/SAR	Y

	Outcomes					Publication methods and results								
Study (name, year)	Change in A1c from baseline	% of patients with A1c <7.0%	% of patients with A1c <6.5%	% patients with decrease in A1c > specific value (specify)	Other A1c outcome	Change in primary and/or secondary outcome with respect to A1c between methods and results	Outcome in methods NR in results	Outcome in methods in-adequately reported in results	Outcome in results NR in methods	Change in analyses between methods and results	Subgroups reported in results that were not described in the methods section	Is SOR/ SAR present based on the publica-tion	ORBIT based on publication (judgment 1)	Change in A1c outcome between original and current outcomes listed in registry
Gupta, 2009	Y	N	N	N	N	N	N	N	N	N	N	N	NA, no SOR/SAR	Y
Hamann, 2008	Y	N	N	N	N	N	N	N	N	N	N	N	NA, no SOR/SAR	N
Jadzinsky, 2009	Y	Y	Y	N	N	N	N	N	N	N	N	N	NA, no SOR/SAR	N
Kaku, 2009	Y	N	Y	N	N	N	N	N	N	N	N	N	NA, no SOR/SAR	N
Nauck, 2007 Seck, 2010	Y	Y	Y	N	N	N	N	N	N	N	N	Y	E	N
Nauck, 2009	Y	Y	Y	N	N	N	N	N	N	N	N	Y	A	N
Perez, 2009	Y	Y	N	N	N	N	N	N	N	N	N	N	NA, no SOR/SAR	N
Pratley, 2010	Y	Y	Y	N	N	N	N	N	N	N	Y	Y	SOR/SAR but ORBIT does not apply	N
Raskin, 2009	Y	Y	Y	Y	Y <7.5%	N	N	Y	N	N	N	Y	A	N
Raz, 2008	Y	Y	N	N	N	N	N	N	N	N	N	N	NA, no SOR/SAR	N
Rigby, 2009	Y	Y	N	N	N	N	N	N	N	N	N	N	NA, no SOR/SAR	N
Robbins, 2007	Y	Y	Y	N	N	N	N	N	N	N	Y	Y	SOR/SAR but ORBIT does not apply	N
Rosenstock, 2006	Y	Y	Y	N	N	N	N	N	N	N	N	N	NA, no SOR/SAR	N
Scott, 2008	Y	Y	N	N	N	N	N	N	N	N	Y	Y	D	N
Seino, 2010	Y	Y	Y	N	N	N	N	N	N	N	N	N	NA, no SOR/SAR	N

	Outcomes					Publication methods and results								
Study (name, year)	**Change in A1c from baseline**	**% of patients with A1c <7.0%**	**% of patients with A1c <6.5%**	**% patients with decrease in A1c > specific value (specify)**	**Other A1c outcome**	**Change in primary and/or secondary outcome with respect to A1c between methods and results**	**Outcome in methods NR in results**	**Outcome in methods in-adequately reported in results**	**Outcome in results NR in methods**	**Change in analyses between methods and results**	**Subgroups reported in results that were not described in the methods section**	**Is SOR/ SAR present based on the publica-tion**	**ORBIT based on publication (judgment 1)**	**Change in A1c outcome between original and current outcomes listed in registry**
van der Meer, 2009	Y	N	N	N	N	N	N	N	Y	N	N	Y	SOR/SAR but ORBIT does not apply	NA

Table B1. Oral Hypoglycemic Agents: Summary Table, continued

	Registry without registry results							Results in registry		Comments
Study (name, year)	Change in 1º outcome between the registry and the pub	Change in status of A1c with respect to 1º and 2º outcome between the registry and the pub	Change in follow-up interval between registry and pub	Index outcome missing in the pub but in the registry	Index outcome reported in the pub but NR in the registry	Is SOR and/or SAR present based on both pub and registry? (Not including registry results)	ORBIT based on registry, not including registry results (judgment 2)	Differences between registry results and pub results	ORBIT based on pub and registry, including registry results (judgment 3)	Describe any "Y" responses
Aschner, 2010	N	N	N	N	Y	Y	NA, ORBIT classes don't apply	N	NA	2. Categorical A1c not mentioned in registry methods but reported in publication results
Bakris, 2006	N	N	N	N	Y	Y	NA, ORBIT classes don't apply	NA, no results	NA	2. A1c not mentioned in registry methods but reported in publication results
Bunck, 2009	N	N	N	N	Y	Y	NA, ORBIT classes don't apply	Y	NA, ORBIT classes do not apply	1. Publication didn't adequately specify A1c outcomes 2. A1c not mentioned in registry methods but reported in publication results 3. Publication and registry results differ
Defronzo, 2009	N	N	N	N	Y	N	NA, ORBIT classes don't apply	N	NA	1. Publication reported subgroup results not mentioned in the methods section; Subgroups inadequately reported

	Registry without registry results							Results in registry		Comments
Study (name, year)	**Change in 1º outcome between the registry and the pub**	**Change in status of A1c with respect to 1º and 2º outcome between the registry and the pub**	**Change in follow-up interval between registry and pub**	**Index outcome missing in the pub but in the registry**	**Index outcome reported in the pub but NR in the registry**	**Is SOR and/or SAR present based on both pub and registry? (Not including registry results)**	**ORBIT based on registry, not including registry results (judgment 2)**	**Differences between registry results and pub results**	**ORBIT based on pub and registry, including registry results (judgment 3)**	**Describe any "Y" responses**
Defronzo, 2010	N	N	N	N	Y	Y	NA, ORBIT classes don't apply	Y	NA, ORBIT classes do not apply	1. Exact P-value NR (P<0.05) in publication; "A1c" specified in methods of publication: results reported change in A1c at 20w 2. A1c added to registry (including results) after study completed 3. Publication reports additional between-group P-values compared to registry
Garber, 2009	N	N	N	N	Y	Y	NA, ORBIT classes don't apply	N	NA	1. Subgroups reported in publication but NR in publication methods; NI analysis proposed in methods of publication but NR in publication results or in registry 2. Categorical A1c not mentioned in registry methods but reported in publication results
Goldberg, 2005	N	N	Y	N	N	Y	ORBIT G or H	N	NA	2. Followup interval of 39w mentioned in registry but NR in publication ORBIT based on registry was either G or H: clinical judgment NA to followup interval

	Registry without registry results							Results in registry		Comments
Study (name, year)	**Change in 1º outcome between the registry and the pub**	**Change in status of A1c with respect to 1º and 2º outcome between the registry and the pub**	**Change in follow-up interval between registry and pub**	**Index outcome missing in the pub but in the registry**	**Index outcome reported in the pub but NR in the registry**	**Is SOR and/or SAR present based on both pub and registry? (Not including registry results)**	**ORBIT based on registry, not including registry results (judgment 2)**	**Differences between registry results and pub results**	**ORBIT based on pub and registry, including registry results (judgment 3)**	**Describe any "Y" responses**
Goldstein, 2007 Williams-Herman, 2009	N	N	N	N	Y	Y	NA, ORBIT classes don't apply	N	NA	2. A1c not mentioned in registry methods but reported in publication results
Gupta, 2009	Y	N	N	N	Y	Y	NA, ORBIT classes don't apply	NA, no results	NA	2. A1c not mentioned in registry methods but reported in publication results
Hamann, 2008	N	N	N	N	N	N	NA. nor SOR/SAR	NA, no results	NA	
Jadzinsky, 2009	N	N	N	N	N	N	NA, no SOR/SAR	N	NA	
Kaku, 2009	N	N	N	N	Y	Y	NA, ORBIT classes don't apply	NA, no results	NA	2. Categorical A1c not mentioned in registry methods but reported in publication results
Nauck, 2007 Seck, 2010	N	N	N	N	Y	Y	NA, ORBIT classes don't apply	N	NA	1. A1c <6.5% NR in results in Seck at 2-y F/U, but data were available 2. Categorical A1c not mentioned in registry methods but reported in publication results
Nauck, 2009	N	N	N	N	Y	Y	NA, ORBIT classes don't apply	N	NA	1. Results reported as not significant 2. Categorical A1c not mentioned in registry methods but reported in publication results

	Registry without registry results							Results in registry		Comments
Study (name, year)	Change in 1º outcome between the registry and the pub	Change in status of A1c with respect to 1º and 2º outcome between the registry and the pub	Change in follow-up interval between registry and pub	Index outcome missing in the pub but in the registry	Index outcome reported in the pub but NR in the registry	Is SOR and/or SAR present based on both pub and registry? (Not including registry results)	ORBIT based on registry, not including registry results (judgment 2)	Differences between registry results and pub results	ORBIT based on pub and registry, including registry results (judgment 3)	Describe any "Y" responses
Perez, 2009	N	N	Y	N	Y	Y	NA, ORBIT classes don't apply	N	NA	2. Primary outcome changed in the registry from 24w to 24w or final visit after data collection was complete; Categorical A1c not mentioned in registry methods but reported in publication results; Additional between-group comparisons or monotherapies provided in registry but not publication
Pratley, 2010	N	N	N	N	Y	Y	NA, ORBIT classes don't apply	N	NA	1. Subgroup analysis not mentioned in the methods section but reported in the publication results 2. Categorical outcomes were added to the registry after the study was completed
Raskin, 2009	N	N	N	N	Y	Y	NA, ORBIT classes don't apply	NA, no results	NA	1. Results inadequately reported in the publication 2. Categorical A1c not mentioned in registry methods but reported in publication results ("sudden levels A1c")

	Registry without registry results							Results in registry		Comments
Study (name, year)	**Change in 1º outcome between the registry and the pub**	**Change in status of A1c with respect to 1º and 2º outcome between the registry and the pub**	**Change in follow-up interval between registry and pub**	**Index outcome missing in the pub but in the registry**	**Index outcome reported in the pub but NR in the registry**	**Is SOR and/or SAR present based on both pub and registry? (Not including registry results)**	**ORBIT based on registry, not including registry results (judgment 2)**	**Differences between registry results and pub results**	**ORBIT based on pub and registry, including registry results (judgment 3)**	**Describe any "Y" responses**
Raz, 2008	N	N	N	N	Y	Y	NA, ORBIT classes don't apply	N	NA	2. Categorical A1c not mentioned in registry methods but reported in publication results F/U interval specified in publication and registry as 18w, secondary outcome 30w (specified 4/10) (latter makes more sense clinically)
Rigby, 2009	N	N	N	N	Y	Y	NA, ORBIT classes don't apply	N	NA	2. Categorical A1c not mentioned in registry methods but reported in publication results
Robbins, 2007	N	N	N	N	Y	Y	NA, ORBIT classes don't apply	NA, no results	NA	1. Subgroups reported in results that were not specified in methods 2. Categorical A1c not mentioned in registry methods but reported in publication results; A1c not adequately specified in the registry including not specifying followup interval
Rosenstock, 2006	N	N	N	N	N	N	NA, nor SOR/SAR	NA, no results	NA	

	Registry without registry results							Results in registry		Comments
Study (name, year)	Change in 1º outcome between the registry and the pub	Change in status of A1c with respect to 1º and 2º outcome between the registry and the pub	Change in follow-up interval between registry and pub	Index outcome missing in the pub but in the registry	Index outcome reported in the pub but NR in the registry	Is SOR and/or SAR present based on both pub and registry? (Not including registry results)	ORBIT based on registry, not including registry results (judgment 2)	Differences between registry results and pub results	ORBIT based on pub and registry, including registry results (judgment 3)	Describe any "Y" responses
Scott, 2008	N	N	N	N	Y	Y	NA, ORBIT classes don't apply	N	NA	1. Prespecified subgroups inadequately reported in the results of the publication 2. Categorical A1c not mentioned in registry methods but reported in publication results
Seino, 2010	N	N	N	N	Y	Y	NA, ORBIT classes don't apply	N	NA	2. Categorical A1c not mentioned in registry methods but reported in publication results
van der Meer, 2009	N	N	N	N	Y	Y	NA, ORBIT classes don't apply	NA, no results	NA	1. A1c not adequately specified in the methods section 2. A1c not mentioned in registry methods but reported in publication results

Abbreviations: Y, Yes; N, No NA, not applicable; NR, not reported; pub, publication; 1º, primary; 2º, secondary.

Table B2. Oral Hypoglycemic Agents: Predictors of SOR/SAR

Study (name, year)	Author industry affiliation for predictors analysis	% of authors with pharma affiliation	Study drug made by study sponsor	Was there any assistance authoring the publication provided by study funder?	Study was registered in Clinicaltrials.gov	Timing of trial registration with respect to first subject recruitment (before or same month (Y) vs. after or unclear (N))	Registry reports results	Change in A1c measure between original and current registry outcomes
Aschner, 2010	Y	6/7 (85%)	Y	Y	Y	Y	Y	N
Bakris, 2006	Y	5/7 (71%)	Y	N	Y	N	N	N
Bunck, 2009	Y	5/12 (42%)	Y	N	Y	N	Y	Y
Defronzo, 2009	Y	3/7 (43%)	Y	N	Y	Y	Y	N
Defronzo, 2010	Y	4/6 (67%)	Y	N	Y	Y	Y	Y
Garber, 2009	Y	2/9 (22%)	Y	Y	Y	Y	Y	N
Goldberg, 2005	Y	6/10 (60%)	Y	N	Y	N	N	N
Goldstein, 2007 Williams-Herman, 2009	Y	Goldstein: 3/5 (60%) Williams-Herman: 8/8 (100%)	Y	Y	Y	Y	Y	Y
Gupta, 2009	N	0/4 (0%)	Y	N	Y	N	N	N
Hamann, 2008	Y	3/5 (60%)	Y	N	Y	N	N	N
Jadzinsky, 2009	Y	3/6 (50%)	Y	N	Y	Y	Y	N
Kaku, 2009	N	0/1 (0%)	Y	Y	N	N	N	Y
Nauck, 2007 Seck, 2010	Y	Nauck: 4/5 (80%) Seck: 7/8 (86%)	Y	Y	Y	N	Y	N
Nauck, 2009	Y	2/9 (22%)	U	Y	Y	Y	Y	N
Perez, 2009	Y	3/4 (75%)	Y	Y	Y	N	Y	N
Pratley, 2010	Y	2/9 (22%)	Y	Y	Y	Y	Y	Y
Raskin, 2009	Y	2/4 (50%)	Y	Y	Y	Y	N	N
Raz, 2008	Y	8/9 (89%)	Y	N	Y	Y	Y	N
Rigby, 2009	Y	4/6 (66%)	Y	Y	Y	N	Y	N
Robbins, 2007	Y	5/10 (50%)	Y	N	Y	N	N	N
Rosenstock, 2006	Y	4/6 (66%)	Y	Y	Y	N	N	N
Scott, 2008	Y	3/4 (75%)	Y	N	Y	N	Y	N
Seino, 2010	Y	2/4 (50%)	Y	N	Y	Y	Y	N
van der Meer, 2009	Y	1/14 (7%)	Y	N	Y	N	N	N

Abbreviations: Y, Yes; N, No; U, Unsure

Table B3. Osteoporosis: Summary Table

	Outcomes						Publication methods and results							
Study (name, year)	Total fractures	Hip Fractures	Vertebral fractures	Non-vertebral	Wrist	Other	Change in primary and/or secondary outcome with respect to fracture between methods and results	Outcome in methods NR in results	Outcome in methods inadequately reported in results	Outcome in results NR in methods	Change in analyses between methods and results	Subgroups reported in results that were not described in the methods section	Is SOR/SAR present based on the publication	ORBIT based on publication
Barrett-Connor, 2006	Y	N	Y	Y	N	Y	N	N	N	N	N	N	N	NA. no SOR/SAR
Black, 2007	Y	Y	Y	Y	N	N	N	N	N	N	N	N	N	NA. no SOR/SAR
Bonnick, 2006 Companion to Rosen 2005	Y	N	N	N	N	N	N	N	N	Y	N	N	Y	ORBIT NA
Grant, 2005	Y	Y	Y	Y	Y	Y	N	N	N	N	N	N	N	NA. no SOR/SAR
Greenspan, 2006	Y	N	N	N	N	N	N	N	N	Y	N	N	Y	ORBIT NA
Jackson, 2006	Y	N	N	N	N	N	N	N	N	N	N	N	N	NA. no SOR/SAR
McClung, 2006	Y	N	Y	Y	Y	Y	N	N	N	Y	N	N	Y	ORBIT NA
Porthouse, 2005	Y	Y	N	N	N	N	N	N	N	N	N	N	N	NA. no SOR/SAR
Prince, 2006	Y	N	N	N	N	Y	N	N	N	N	N	N	N	NA. no SOR/SAR
Reid, 2006	Y	Y	Y	Y	Y	Y	N	N	N	N	N	N	N	NA. no SOR/SAR

	Outcomes						Publication methods and results							
Study (name, year)	Total fractures	Hip Fractures	Vertebral fractures	Non-vertebral	Wrist	Other	Change in primary and/or secondary outcome with respect to fracture between methods and results	Outcome in methods NR in results	Outcome in methods inadequately reported in results	Outcome in results NR in methods	Change in analyses between methods and results	Subgroups reported in results that were not described in the methods section	Is SOR/SAR present based on the publication	ORBIT based on publication
Rosen, 2005 Companion to Bonnick 2006 but has separate NCT number	Y	N	N	N	N	N	N	N	N	Y	N	N	Y	ORBIT NA
Vogel, 2006	N	Y	Y	N	Y	N	N	N	N	N	N	N	N	NA. no SOR/SAR

	Registry without registry results								Results in registry		Comments
Study (name, year)	**Change in fracture outcome between original and current outcomes listed in registry**	**Change in primary outcome between the registry and the publication**	**Change in status of fractures with respect to primary and secondary outcome between the registry and the publication**	**Change in followup interval between registry and publication**	**Index outcome missing in the publication but in the registry**	**Index outcome reported in the publication but NR in the registry**	**Is SOR and/or SAR present based on both publication and registry? (Not including registry results)**	**ORBIT based on registry, not including registry results**	**Differences between registry results and publication results**	**ORBIT based on publication and registry, including registry results**	
Barrett-Connor, 2006	N	N	N	N	N	N	N	NA, no SOR/SAR	NA, no results	NA	
Black, 2007	N	N	N	Y	N	N	N	NA, no SOR/SAR	NA, no results	NA	2. Followup interval not prespecified in the registry
Bonnick, 2006 Companion to Rosen 2005	N	N	N	N	N	Y	Y	NA, ORBIT classes don't apply	NA, no results	NA	1. Fractures were reported in the safety outcomes in the results, but no mention in the methods section (either efficacy or safety) 2. Fractures were presented as an adverse event in the publication results, but not mentioned in the registry
Grant, 2005	N	N	N	Y	N	N	N	NA, no SOR/SAR	NA, no results	NA	2. Followup interval not prespecified in the registry

	Registry without registry results								Results in registry		Comments
Study (name, year)	**Change in fracture outcome between original and current outcomes listed in registry**	**Change in primary outcome between the registry and the publication**	**Change in status of fractures with respect to primary and secondary outcome between the registry and the publication**	**Change in followup interval between registry and publication**	**Index outcome missing in the publication but in the registry**	**Index outcome reported in the publication but NR in the registry**	**Is SOR and/or SAR present based on both publication and registry? (Not including registry results)**	**ORBIT based on registry, not including registry results**	**Differences between registry results and publication results**	**ORBIT based on publication and registry, including registry results**	
Greenspan, 2006	N	N	N	N	N	Y	Y	NA, ORBIT classes don't apply	NA, no results	NA	1. Fractures not mentioned in the publication methods but reported as a safety outcome 2. Fractures not mentioned in the registry methods but reported as a safety outcome in the publication
Jackson, 2006	N	N	N	N	N	Y	Y	NA, ORBIT classes don't apply	NA, no results	NA	2. Registry does not specify any outcomes

	Registry without registry results								Results in registry		Comments
Study (name, year)	**Change in fracture outcome between original and current outcomes listed in registry**	**Change in primary outcome between the registry and the publication**	**Change in status of fractures with respect to primary and secondary outcome between the registry and the publication**	**Change in followup interval between registry and publication**	**Index outcome missing in the publication but in the registry**	**Index outcome reported in the publication but NR in the registry**	**Is SOR and/or SAR present based on both publication and registry? (Not including registry results)**	**ORBIT based on registry, not including registry results**	**Differences between registry results and publication results**	**ORBIT based on publication and registry, including registry results**	
McClung, 2006	N	N	N	N	N	Y	Y	NA, ORBIT classes don't apply	Y	ORBIT NA	1. Fractures not mentioned in the publication methods but reported as a safety outcome 2. Fractures not mentioned in the registry methods but reported as a safety outcome in the publication 3. Fractures not reported in the registry results but reported as a safety outcome in the publication

	Registry without registry results								Results in registry		Comments
Study (name, year)	**Change in fracture outcome between original and current outcomes listed in registry**	**Change in primary outcome between the registry and the publication**	**Change in status of fractures with respect to primary and secondary outcome between the registry and the publication**	**Change in followup interval between registry and publication**	**Index outcome missing in the publication but in the registry**	**Index outcome reported in the publication but NR in the registry**	**Is SOR and/or SAR present based on both publication and registry? (Not including registry results)**	**ORBIT based on registry, not including registry results**	**Differences between registry results and publication results**	**ORBIT based on publication and registry, including registry results**	
Porthouse, 2005	N	N	N	Y	N	Y	Y	NA, ORBIT classes don't apply	NA, no results	NA	2. Secondary outcomes of hip fractures and hip-wrist fractures were added to the publication but not mentioned in the registry methods; followup not mentioned in the registry
Prince, 2006	N	N	N	N	N	Y	Y	NA, ORBIT classes don't apply	NA, no results	NA	2. Fractures not mentioned in the registry methods but reported in the publication results
Reid, 2006	N	N	N	N	N	N	N	NA, no SOR/SAR	NA, no results	NA	

	Registry without registry results								Results in registry		Comments
Study (name, year)	**Change in fracture outcome between original and current outcomes listed in registry**	**Change in primary outcome between the registry and the publication**	**Change in status of fractures with respect to primary and secondary outcome between the registry and the publication**	**Change in followup interval between registry and publication**	**Index outcome missing in the publication but in the registry**	**Index outcome reported in the publication but NR in the registry**	**Is SOR and/or SAR present based on both publication and registry? (Not including registry results)**	**ORBIT based on registry, not including registry results**	**Differences between registry results and publication results**	**ORBIT based on publication and registry, including registry results**	
Rosen, 2005 Companion to Bonnick 2006 but has separate NCT number	N	N	N	N	N	Y	Y	NA, ORBIT classes don't apply	NA, no results	NA	1. Fractures were reported in the safety outcomes in the results, but no mention in the methods section (either efficacy or safety) 2. Fractures were presented as an adverse event in the publication results, but not mentioned in the registry
Vogel, 2006	N	N	N	N	N	N	N	NA, no SOR/SAR	NA, no results	NA	

Abbreviations: Y, Yes; N, No; U, Unsure; CI, confidence interval; OR, odds ratio.

Table B4. Osteoporosis: Predictors of SOR/SAR

	Predictors						
Study (name, year)	**Author industry affiliation**	**Percent of authors with pharma affiliation**	**Study drug made by study sponsor**	**Was there any assistance authoring the publication provided by study funder?**	**Study was registered in Clinicaltrials.gov**	**Timing of trial registration with respect to first subject recruitment (before or same month (Y) vs. after or unclear (N))**	**Registry reports results**
Barrett-Connor, 2006	Y	2/8 (25%)	Y	N	Y	N	N
Black, 2007	Y	7/21 (33%)	Y	Y	Y	N	N
Bonnick, 2006 Companion to Rosen 2005	Y	3/11 (27%)	Y	N	Y	N	N
Grant, 2005	N	0	N	N	N	N	N
Greenspan, 2006	N	0/5 (0%)	N	N	Y	N	N
Jackson, 2006	N	0/47 (0%)	N	N	Y	U (study start date NR in registry	N
McClung, 2006	Y	4/16 (25%)	Y	Y	Y	N	Y
Porthouse, 2005	N	0/15 (0%)	N	N	N	N	N
Prince, 2006	N	0/4 (0%)	N	N	N	N	N
Reid, 2006	N	0/8 (0%)	N	N	N	N	N
Rosen, 2005 Companion to Bonnick 2006 but has separate NCT number	N	4/11 (36%)	Y	N	Y	N	N
Vogel, 2006	N	0/21 (0%)	Y	N	Y	N	N

Abbreviations: Y, Yes; N, No; U, Unsure.

Table B5. Oral Hypoglycemic Agents: Data Abstraction: Study Funder and Conflicts of Interest

		Study characteristics, from the publication						
Author, year	**Journal**	**Study funder**	**Percentage of authors employed by pharmaceutical industry**	**Percentage of authors with COI from industry**	**What company makes intervention drug?**	**What company makes comparator drug?**	**Was there any assistance authoring the publication?**	**Did the publication indicate that the trial was registered?**
Aschner 2010	Diabetes, Obesity and Metabolism	Merck & Co., Inc.	6/7 (85%)	1/1 (100%)	Sitagliptin: Merck & Co., Inc	Metformin: generic Fortamet: Shionogi Pharma, Inc.	Yes	Yes
Bakris 2006	Journal of Hypertension	GlaxoSmithKline	5/7 (71%)	1/2 (50%)	Rosiglitazone: GlaxoSmithKline (Avandia) and metformin: generic	Metformin: generic Glyburide: generic and sanofi-aventis U.S. LLC (DiaBeta)	Yes	No
Bunck 2009	Diabetes Care	The study was sponsored by Amylin Pharmaceuticals and Eli Lilly and Company	5/12 (42%)	6/7 (86%)	Exenatide: Amylin and Eli Lilly	Insulin Glargine: sanofi-aventis (Lantus)	NR	Yes
Defronzo 2009	Diabetes Care	Bristol-Myers Squibb and AstraZeneca	3/7 (43%)	4/4 (100%)	Saxagliptin: Bristol-Myers Squibb (Onglyza)	Metformin: generic	Yes	Yes
Defronzo 2010	Diabetes Care	NR	4/6 (67%)	2/2 (100%)	Exenatide: Amylin and Eli Lilly	Rosiglitazone: GlaxoSmithKline	NR	Yes
Garber 2009	The Lancet	Novo Nordisk A/S	2/9 (22%)	4/7 (57%)	Liraglutide: Novo Nordisk (Victoza)	Glimepiride: generic	Yes	Yes
Goldberg 2005	Diabetes Care	Eli Lilly and Takeda Pharmaceuticals	6/10 (60%)	0/4 (0%)	Pioglitazone: Takeda Pharmaceuticals	Rosiglitazone: GlaxoSmithKline (Avandia)	No	No
Goldstein 2007; Williams-Herman 2009	Clinical Care / Education / Nutrition / Psychosocial Research	Merck & Co., Inc.	Goldstein: 3/5 (60%) Williams-Herman: 8/8 (100%)	Goldstein: 2/2 (100%) Williams-Herman: NA	Sitagliptin: Merck and Co., Inc.	Metformin: generic	Goldstein: Yes Williams-Herman: NR	Yes
Gupta 2009	Diabetes, Obesity and Metabolism	Takeda Pharmaceuticals	0/4 (0%)	2/4 (50%)	Pioglitazone: Takeda Pharmaceuticals	Metformin: generic	NR	Yes
Hamann 2008	Exp Clin Endocrinol Diabetes	NR (clearly GlaxoSmithKline)	3/5 (60%)	0/2 (0%)	Rosiglitazone: GlaxoSmithKline (Avandia)	Metformin: generic	Yes	No

		Study characteristics, from the publication						
Author, year	**Journal**	**Study funder**	**Percentage of authors employed by pharmaceutical industry**	**Percentage of authors with COI from industry**	**What company makes intervention drug?**	**What company makes comparator drug?**	**Was there any assistance authoring the publication?**	**Did the publication indicate that the trial was registered?**
Jadzinsky 2009	Diabetes, Obesity and Metabolism	Bristol-Myers Squibb	3/6 (50%)	3/3 (100%)	Saxagliptin: Bristol-Myers Squibb (Onglyza)	Metformin: generic	Yes	Yes
Kaku 2009	Current Med Res and opinion	Takeda Pharmaceutical Co	0/1 (0%)	0/1 (0%)	Combination therapy: Pioglitazone: Takeda Pharmaceuticals; Metformin: generic	Metformin: generic	Yes	Yes
Nauck 2007; Seck 2010	Diabetes, Obesity and Metabolism	Nauck: Merck & Co., Inc. Seck: Merck & Co., Inc.	Nauck: 4/5 (80%) Seck: 7/8 (86%)	Nauck: 1/1 (100%) Seck: 1/1 (100%)	Nauck: sitagliptin: Merck Seck: sitagliptin: Merck	Nauck: glipizide: generic Seck: glipizide: generic	Nauck: Yes Seck: No	Nauck: No Seck: Yes
Nauck 2009	Diabetes Care	NR	2/9 (22%)	2/7 (29%)	Liraglutide: Novo Nordisk	Glimepiride: generic Amaryl: sanofi-aventis Metformin: generic	Yes	Yes
Perez 2009	Current Med Res and opinion	Takeda Global Research and Development Center	3/4 (75%)	1/1 (100%)	Pioglitazone: Takeda Pharmaceuticals	Metformin: generic	Yes	Yes
Pratley 2010	Lancet	NovoNordisk	2/9 (22%)	7/7 (100%)	Liraglutide: Novo Nordisk Pharmaceuticals Sitagliptin: Merck & Co	Metformin: generic	Yes	Yes
Raskin 2009	Diabetes, Obesity and Metabolism	Novo Nordisk	2/4 (50%)	2/2 (100%)	Repaglinide: Novo Nordisk (Prandin) Metformin: generic	Rosiglitazone: GlaxoSmithKline Metformin: generic	Yes	No
Raz 2008	Current Med Res and opinion	Merck & Co, USA	8/9 (89%)	0/1 (0%)	Sitagliptin: Merck & Co., Inc	Metformin: generic	NR	No

		Study characteristics, from the publication						
Author, year	**Journal**	**Study funder**	**Percentage of authors employed by pharmaceutical industry**	**Percentage of authors with COI from industry**	**What company makes intervention drug?**	**What company makes comparator drug?**	**Was there any assistance authoring the publication?**	**Did the publication indicate that the trial was registered?**
Rigby 2009	Endocrine Practice	Daiichi Sankyo, Inc.	4/6 (66%)	1/2 (50%)	Colesevelam: Daiichi Sankyo (Welcho); Rosiglitazone: GlaxoSmithKline (Avandia); Sitagliptin: Merck & Co., Inc.	Metformin: generic	Yes	No
Robbins 2007	Clinical Therapeutics	NR (clearly Eli Lilly)	5/10 (50%)	3/5 (60%)	Insulin lispro protamine: Eli Lilly Metformin: generic	Insulin glargine: sanofi-aventis Metformin: generic	No	Yes
Rosenstock 2006	Diabetes, Obesity and Metabolism	GlaxoSmithKline	4/6 (66%)	0/2 (0%)	Rosiglitazone/ metformin combination therapy: GlaxoSmithKline (Avandamet)	Rosiglitazone: GlaxoSmithKline (Avandia) Metformin: generic	Yes	No
Scott 2008	Diabetes, Obesity and Metabolism	Merck & Co., Inc.	3/4 (75%)	0/1 (0%)	Sitagliptin: Merck & Co., Inc. Rosiglitazone: GlaxoSmithKline	Metformin: generic	NR	Yes
Seino 2010	Current Med Res and opinion	Novo Nordisk Pharma Ltd, Japan	2/4 (50%)	2/2 (100%)	Liraglutide: Novo Nordisk (Victoza)	Glibenclamide: Taisho Pharmaceutical Co.	Yes	Yes

		Study characteristics, from the publication						
Author, year	**Journal**	**Study funder**	**Percentage of authors employed by pharmaceutical industry**	**Percentage of authors with COI from industry**	**What company makes intervention drug?**	**What company makes comparator drug?**	**Was there any assistance authoring the publication?**	**Did the publication indicate that the trial was registered?**
van der Meer 2009	Circulation	This investigator-initiated study was supported by Eli Lilly, the Netherlands, which has a partnership with Takeda, the manufacturer of pioglitazone. Metformin tablets and matching placebos were kindly provided by Merck, the Netherlands.	1/14 (7%)	2/13 (15%)	Pioglitazone: Takeda Pharmaceuticals (Actos)	Metformin: generic	NR	Yes

Table B6. Oral Hypoglycemic Agents: Data Abstraction, Study Characteristics

	Publication, methods section								
Author, year	**Study design**	**Intervention**	**Comparison**	**Total N**	**Primary outcome stated in the study (relevant to index outcome)**	**Relevant secondary outcomes (relevant to index outcome)**	**Followup intervals (weeks)** **F/U 1** **F/U 2** **F/U 3**	**Analysis set (definition from study)** **Definition of analysis set**	**How handled missing values**
Aschner 2010	Non-inferiority, parallel group	Sitagliptin	Metformin	1050	HbA1c change from baseline at week 24	Proportions of patients with HbA1c <7 or<6.5%	24 NA NA	PP; Secondary analyses was FAS (all randomized patients with 1+ study drug dose and B/ and 1+ F/U measures) Patients who completed the study and did not have any reasons for exclusion from this population, including absence of baseline or on-treatment data at the week 24 visit or major protocol violations (e.g. drug compliance <75%, addition of non-study antihypergly-cemic agent or incorrect double-blind study medication). Also could be excluded for lack of efficacy.	LOCF

	Publication, methods section								
Author, year	**Study design**	**Intervention**	**Comparison**	**Total N**	**Primary outcome stated in the study (relevant to index outcome)**	**Relevant secondary outcomes (relevant to index outcome)**	**Followup intervals (weeks)** **F/U 1** **F/U 2** **F/U 3**	**Analysis set (definition from study)** **Definition of analysis set**	**How handled missing values**
Bakris 2006	Parallel group, double-blind, superiority	Rosiglitazone plus metformin	Glyburide plus metformin	389	None	Additional pharmacodynamic end points included change from baseline at week 32 in HbA1c	32 NA NA	ITT All randomized patients who had at least one postbaseline data point for any efficacy parameter; for the secondary population (completers) ITT population with no use of LOCF	LOCF
Bunck 2009	Parallel group, open label	Exenatide plus metformin	Insulin glargine plus metformin	69	None	Glycemic control	52 64 (for A1c and body weight) NA	NR in methods section; ITT listed in flow diagram NR	NR

	Publication, methods section								
Author, year	**Study design**	**Intervention**	**Comparison**	**Total N**	**Primary outcome stated in the study (relevant to index outcome)**	**Relevant secondary outcomes (relevant to index outcome)**	**Followup intervals (weeks)** **F/U 1** **F/U 2** **F/U 3**	**Analysis set (definition from study)** **Definition of analysis set**	**How handled missing values**
Defronzo 2009	Parallel group trial	Saxagliptin plus metformin	Metformin plus placebo	743	Change from baseline in A1C to week 24	Percentage of patients at the glycemic target (defined as A1C <7.0%)	24 42-month long term extension NA	Efficacy analyses were performed on the randomly assigned patient population Consisting of randomly assigned patients who received at least one dose of study medication and had a baseline and at least one postbaseline measurement	LOCF
Defronzo 2010	Parallel group, open label	Exenatide injection Rosiglitazone	Combination of exenatide plus rosiglitazone	137	None	Efficacy measurements included A1C, glucose, insulin, C-peptide, lipids, and body weight	20 NA NA	ITT Included participants with a baseline and at least one post baseline value	NR
Garber 2009	Parallel group trial; superiority; non-inferiority analysis mentioned but not presented	Subcutaneous liraglutide	Oral glimepiride	746	Change in value of HbA1c from baseline to 52 weeks	Proportion of patients achieving A1c <7.0% and >6.5%	52 NA NA	ITT Participants exposed to at least one dose	LOCF

	Publication, methods section								
Author, year	**Study design**	**Intervention**	**Comparison**	**Total N**	**Primary outcome stated in the study (relevant to index outcome)**	**Relevant secondary outcomes (relevant to index outcome)**	**Followup intervals (weeks)** **F/U 1** **F/U 2** **F/U 3**	**Analysis set (definition from study)** **Definition of analysis set**	**How handled missing values**
Goldberg 2005	Parallel group trial	Pioglitazone	Rosiglitazone	802	None	A1C: mentioned in analysis section but not in prior parts of methods section	24 NA NA	Infer ITT (see definition) Efficacy analyses were conducted on subjects providing a baseline measurement and at least one postbaseline measurement	LOCF

	Publication, methods section								
Author, year	**Study design**	**Intervention**	**Comparison**	**Total N**	**Primary outcome stated in the study (relevant to index outcome)**	**Relevant secondary outcomes (relevant to index outcome)**	**Followup intervals (weeks)** **F/U 1** **F/U 2** **F/U 3**	**Analysis set (definition from study)** **Definition of analysis set**	**How handled missing values**
Goldstein 2007; Williams-Herman 2009	Parallel group trial	Sitagliptin/metformin	Placebo	1091	Change from baseline at week 24 was assessed for A1C	Proportion <7.0 and <6.5% in each RX group Williams-Herman: change from baseline at week 54 for A1c (and others); also mention proportion with A1c <7.0 at week 54 and at both weeks 24 and 54	24 Williams-Herman: 54 NA	Efficacy analyses were based on the APT population Williams: continuation APT (baseline measure, no rescue therapy, 1+ dose study medication, 1+ efficacy measure weeks 24 to 54) All randomized patients who received at least one dose of study treatment and who had both a baseline and at least one postbaseline measurement	LOCF Williams: rescue therapy patients were treated as missing with LOCF
Gupta 2009	Parallel group trial	Pioglitazone plus placebo	Pioglitazone plus ADA diet Metformin plus ADA diet	51	None	Change in A1c	16 NA NA	NR NR explicitly, but reports are on completers only	NR

	Publication, methods section								
Author, year	**Study design**	**Intervention**	**Comparison**	**Total N**	**Primary outcome stated in the study (relevant to index outcome)**	**Relevant secondary outcomes (relevant to index outcome)**	**Followup intervals (weeks)** **F/U 1** **F/U 2** **F/U 3**	**Analysis set (definition from study)** **Definition of analysis set**	**How handled missing values**
Hamann 2008	Parallel group, non-inferiority	Rosiglitazone plus metformin	Sulphonylurea (glibenclamide or gliclazide) Sulphonylurea plus metformin	596	Change in HbA 1c from baseline after 52 weeks of treatment	NA	52 NA NA	ITT ITT without LOCF for A1c, biomarkers, and health outcomes; ITT with LOCF for all other outcomes All randomized subjects who received at least one dose of study medication, had a baseline assessment and at least one corresponding on-therapy assessment for HbA1c	LOCF
Jadzinsky 2009	Parallel group trial	Saxagliptin plus metformin Saxagliptin plus placebo	Metformin plus placebo (metformin)	1306	HbA1c change from baseline to week 24	Proportion of patients achieving HbA1c <7.0% and <6.5%	24 NA NA	All randomized patients who took 1+ dose of study medication	LOCF

	Publication, methods section								
Author, year	**Study design**	**Intervention**	**Comparison**	**Total N**	**Primary outcome stated in the study (relevant to index outcome)**	**Relevant secondary outcomes (relevant to index outcome)**	**Followup intervals (weeks)** **F/U 1** **F/U 2** **F/U 3**	**Analysis set (definition from study)** **Definition of analysis set**	**How handled missing values**
Kaku 2009	Parallel group	Metformin plus pioglitazone	Metformin plus placebo	169	Change in end-of-treatment HbA1c in the FAS population	Secondary endpoints included time course for HbA1c and FBG, and the percentage of patients achieving an HbA1c <6.5%	28 NA NA	FAS A FAS assessment of efficacy was performed in patients receiving >=1 dose of pioglitazone	NR

	Publication, methods section								
Author, year	**Study design**	**Intervention**	**Comparison**	**Total N**	**Primary outcome stated in the study (relevant to index outcome)**	**Relevant secondary outcomes (relevant to index outcome)**	**Followup intervals (weeks)** **F/U 1** **F/U 2** **F/U 3**	**Analysis set (definition from study)** **Definition of analysis set**	**How handled missing values**
Nauck 2007; Seck 2010	Parallel group, non-inferiority Seck: In methods section is stated to be a non-inferiority study at 1 year, with 2 year results having "no predefined efficacy hypotheses"; results presented as superiority	Sitagliptin plus metformin	Glipizide plus metformin	1172	HbA1c change from baseline at week 52	Nauck: Percent < 7.0 and <6.5% Seck: A1c <7.0 at 2y, and <7.0% at both 1 and 2 year	Nauck: 52 Seck: 104 NA	Per-protocol approach Secondary analysis based on all patients treated, with missing values imputed with LOCF Seck: 2 years are PP for efficacy outcome (not non-inferiority) Patients who completed all 52 weeks of treatment and did not have any reasons for exclusion from this population, including no baseline data, no treatment data at Week 52 or major protocol violations	LOCF for APT analyses

	Publication, methods section								
Author, year	**Study design**	**Intervention**	**Comparison**	**Total N**	**Primary outcome stated in the study (relevant to index outcome)**	**Relevant secondary outcomes (relevant to index outcome)**	**Followup intervals (weeks)** **F/U 1** **F/U 2** **F/U 3**	**Analysis set (definition from study)** **Definition of analysis set**	**How handled missing values**
Nauck 2009	Parallel group trial; both a superiority and NI trial (liraglutide and metformin is significantly better or at least as good as metformin)	Subcutaneous liraglutide Glimepiride	Placebo	1091	Change in A1C at the end of the study (26 weeks)	None explicitly listed, but in statistical section, percentage with A1c < 0.7% and <=0.6.5%	26 NA NA	ITT Subjects who were exposed to at least one dose of trial product and had one postbaseline measurement of the parameter	LOCF
Perez 2009	Parallel group	Pioglitazone/ metformin combination therapy	Pioglitazone mono therapy, metformin mono therapy	600	Change in HbA1c from baseline to final visit or early termination	Percent with A1c <=7%; changes from baseline to week 24 (or early termination)	24 NA NA	FAS >=1 dose drug, baseline, and at least one treatment value	LOCF from last post-baseline measurement

	Publication, methods section								
Author, year	**Study design**	**Intervention**	**Comparison**	**Total N**	**Primary outcome stated in the study (relevant to index outcome)**	**Relevant secondary outcomes (relevant to index outcome)**	**Followup intervals (weeks)** **F/U 1** **F/U 2** **F/U 3**	**Analysis set (definition from study)** **Definition of analysis set**	**How handled missing values**
Pratley 2010	Parallel group, open label, non-inferiority followed by superiority	Subcutaneous liraglutide	Oral sitagliptin	665	Change in HbA1c from baseline to week 26	Proportions of participants reaching HbA1c targets of less than 7·0% or of 6·5% or lower; and a composite endpoint of proportions of participants with HbA1c of less than 7·0%, with no hypoglycemia	26 NA NA	NI: Full analysis set and per protocol sets: superiority: FAS: secondary analyses on the FAS FAS: randomized participants who were exposed to at least one dose of trial drug and with at least one HbA1c measurement taken after baseline	LOCF
Raskin 2009	Parallel group trial; 2 non-inferiority comparisons	Repaglinide/ metformin	Rosiglitazone/ metformin	561	HbA1c change from baseline	Percentage of subjects A1c <7.0, 7.5, 6.5%	26 NA NA	ITT Those randomized subjects who received at least one dose of trial medication and had at least one postbaseline assessment	LOCF

	Publication, methods section								
Author, year	**Study design**	**Intervention**	**Comparison**	**Total N**	**Primary outcome stated in the study (relevant to index outcome)**	**Relevant secondary outcomes (relevant to index outcome)**	**Followup intervals (weeks)** **F/U 1** **F/U 2** **F/U 3**	**Analysis set (definition from study)** **Definition of analysis set**	**How handled missing values**
Raz 2008	Parallel group	Sitagliptin plus metformin	Metformin plus placebo	190	Reduction in A1c at 18 weeks	30 week A1c; percent of patients reaching goal A1c <7.0%	18 30 NA	FAS FAS= all randomized with >=1 dose and baseline plus 1 F/U measure at week 6	LOCF from start of rescue RX; LOCF for missing data
Rigby 2009	Parallel group, open label, superiority	Colesevelam	Rosiglitazone, sitagliptin	169	Change in A1C from baseline to week 16	Change in A1C from baseline to Week 8. Percentage of subjects who achieved an A1c reduction of ≥0.7% and <7.0% at 16 weeks. %Percentage of subjects who achieved A1c target of <7.0% at 16 weeks	16 8 NA	FAS All randomized subjects who had taken ≥1 dose of study medication and had a baseline and ≥1 post-baseline A1C measurement	LOCF
Robbins 2007	Parallel group, open label	Insulin lispro plus metformin	Insulin glargine HS plus metformin	317	HbA1c at endpoint	NA	12 24 NA	ITT Analyses were performed on data from randomized patients who received ≥1 dose of study drug	LOCF

	Publication, methods section								
Author, year	**Study design**	**Intervention**	**Comparison**	**Total N**	**Primary outcome stated in the study (relevant to index outcome)**	**Relevant secondary outcomes (relevant to index outcome)**	**Followup intervals (weeks)** **F/U 1** **F/U 2** **F/U 3**	**Analysis set (definition from study)** **Definition of analysis set**	**How handled missing values**
Rosenstock 2006	Double-blind, parallel group; superiority	Rosiglitazone plus metformin	Rosiglitazone or metformin	468	A1c from baseline to week 32	The proportions of patients achieving recommended A1c targets (<7.0 and <6.5%)	32 NA NA	ITT All randomized patients who received at least one dose of study medication and who had at least one valid on-therapy observation for an efficacy variable	LOCF
Scott 2008	Parallel group, superiority (of sitagliptin versus placebo)	Metformin plus sitagliptin or metformin plus rosiglitazone	Metformin plus placebo	273	Change in HbA1c from baseline	Proportion of patients achieving HbA1c < 7%	18 NA NA	Efficacy analyses were based on the APT population All randomized patients who received at least one dose of study drug and who had both a baseline and at least one postbaseline measurement	LOCF

	Publication, methods section								
Author, year	**Study design**	**Intervention**	**Comparison**	**Total N**	**Primary outcome stated in the study (relevant to index outcome)**	**Relevant secondary outcomes (relevant to index outcome)**	**Followup intervals (weeks)** **F/U 1** **F/U 2** **F/U 3**	**Analysis set (definition from study)** **Definition of analysis set**	**How handled missing values**
Seino 2010	Parallel group, double dummy, non-inferiority; superiority	Liraglutide plus placebo	Glibenclamide plus placebo	411	A1c at 24 weeks	Percent with A1c <7.0 ("post hoc") or <6.5%	24 Open-label extension to week 52 described but results NR herein NA	FAS FAS= >=1 dose drug	LOCF
van der Meer 2009	Parallel group, double dummy, superiority	Pioglitazone	Metformin	78	None	A1c	24 NA NA	NR NR	NR

Table B7. Oral Hypoglycemic Agents: Data Abstraction, Study Registration Information

	Clinicaltrials.gov registration (or other registry) information									
Author, year	**Subgroups specified in the methods section**	**Registry Registry number**	**Study sponsor (as noted in registry)**	**First received date**	**Last updated**	**Study start date**	**Study completion date**	**Primary completion date (under tracking in clinicaltrials.gov)**	**Results first received**	**Primary completion date for recruitment**
Aschner 2010	Prespecified: baseline A1c, gender, age, ethnicity, baseline BMI, duration DM, geographic region	Clinicaltrials.gov NCT00449930	Merck	03/19/07	04/20/10	03/2007	07/2008	07/2008	04/23/09	July 2008
Bakris 2006	NR	Clinicaltrials.gov NCT00500955	GlaxoSmithKline	07/12/07	10/01/10	04/2000	06/2004	06/2004	No study results posted	06/2004
Bunck 2009	None	Clinicaltrials.gov NCT00097500	Amylin Pharmaceuticals, Inc.	11/24/04	12/24/10	09/2004	12/2009	12/2009	12/24/10	12/2009
Defronzo 2009	None	Clinicaltrials.gov NCT00121667	Bristol-Myers Squibb	07/18/05	08/05/11	08/2005	02/2010	10/2006	03/15/11	10/2006
Defronzo 2010	None	Clinicaltrials.gov NCT00135330	Amylin; Collaborator: Eli Lilly	08/24/05	07/21/09	10/2005	NR	07/2008	07/21/09	07/2008
Garber 2009	None: subgroups were presented in the results: prior DM treatment	Clinicaltrials.gov NCT00294723- was terminated	Novo Nordisk	02/20/06	03/24/11	02/2006	03/2010	November 2007	02/23/10	11/2008

	Clinicaltrials.gov registration (or other registry) information									
Author, year	**Subgroups specified in the methods section**	**Registry Registry number**	**Study sponsor (as noted in registry)**	**First received date**	**Last updated**	**Study start date**	**Study completion date**	**Primary completion date (under tracking in clinicaltrials.gov)**	**Results first received**	**Primary completion date for recruitment**
Goldberg 2005	Data to be stratified on prior treatment and sex	Clinicaltrials.gov NCT00331487	Takeda Global Research & Development Center, Inc.; Collaborator: Eli Lilly	05/30/06	07/01/10	09/2000	03/2004	03/2004	No study results posted	03/2004
Goldstein 2007; Williams-Herman 2009	OHA status, baseline A1c, sex, age, race, baseline BMI, duration DM, HOMA	Clinicaltrials.gov NCT00103857	Merck	02/15/05	04/07/10	03/2005	02/08	July 2006	2/19/2009	February, 2008
Gupta 2009	None	Clinicaltrials.gov NCT00219440	Pennington Biomedical Research Center	09/14/05	02/02/10	02/2003	12/06	12/06	"No study results posted"	12/2006
Hamann 2008	None	Clinicaltrials.gov NCT00359112	GlaxoSmithKline	07/28/06	05/15/09	02/2004	NR	NR	No study results posted	NR
Jadzinsky 2009	None; "subgroup analyses for baseline HbA1c were prespecified"	Clinicaltrials.gov NCT00327015	Bristol-Myers Squibb	05/15/06	08/04/10	05/2006	12/2008	November 2007	08/17/09	11/07
Kaku 2009	Gender, BMI, pre-treatment of A1c	UMIN-CTR Search Clinical Trials UMIN000001110	Takeda Pharmaceutical Company Limited	04/04/08	08/27/10	04/2005	10/2006	NR	08/27/10; but can't find results	NR

	Clinicaltrials.gov registration (or other registry) information									
Author, year	**Subgroups specified in the methods section**	**Registry Registry number**	**Study sponsor (as noted in registry)**	**First received date**	**Last updated**	**Study start date**	**Study completion date**	**Primary completion date (under tracking in clinicaltrials.gov)**	**Results first received**	**Primary completion date for recruitment**
Nauck 2007; Seck 2010	Subgroups based on baseline A1c	Clinicaltrials.gov NCT00094770	Merck	10/22/04	04/07/10	09/2004	NR	05/2006	09/24/09	05/2006
Nauck 2009	None	Clinicaltrials.gov NCT00318461 NCT00318422 is LEAD-1 study Other study ID: NN2211-1572	Novo Nordisk	04/25/06	04/16/10	05/2006	11/2008	05/07	02/23/10	11/2008
Perez 2009	None	Clinicaltrials.gov NCT00727857	Takeda Global Research & Development Center, Inc.	07/30/08	07/27/11	06/2007	NR	August 2008	8/28/2009	August 2008
Pratley 2010	None, but there is a subgroup reported in the results section of participants with a baseline HbA1c of 9.0% or higher	Clinicaltrials.gov NCT00700817	Novo Nordisk	06/18/08	09/22/11	06/2008	06/2010	06/2009	06/11/10	06/2009

	Clinicaltrials.gov registration (or other registry) information									
Author, year	**Subgroups specified in the methods section**	**Registry Registry number**	**Study sponsor (as noted in registry)**	**First received date**	**Last updated**	**Study start date**	**Study completion date**	**Primary completion date (under tracking in clinicaltrials.gov)**	**Results first received**	**Primary completion date for recruitment**
Raskin 2009	None	Clinicaltrials.gov NCT00399711	Novo Nordisk	11/14/06	09/22/11	11/2006	11/2007	11/2007	No study results posted	11/2007
Raz 2008	Prespecified based on: age, sex, race, duration DM, Bl A1c, others	Clinicaltrials.gov NCT00337610	Merck	06/14/06	05/27/10	06/2006	08/2007	05/2007	09/24/09	05/2007
Rigby 2009	None	Clinicaltrials.gov NCT00484419	Daiichi Sankyo Inc.	06/07/07	06/17/09	05/2007	04/2008	04/2008	04/29/09	04/2008
Robbins 2007	No subgroup analyses reported in methods section, but they appear in the results (number of daily injections; pre-study use of lipid altering medications)	Clinicaltrials.gov NCT00191464	Eli Lilly	09/12/05	10/12/10	12/2003	09/2005	NR	No study results posted	Not stated
Rosenstock 2006	Baseline A1c, gender, treatment	Clinicaltrials.gov NCT00499707	GlaxoSmithKline	07/09/07	03/17/11	10/2003	NR	NR	No study results posted	NR
Scott 2008	Gender, age, race, baseline BMI, baseline A1c, and known duration of type 2 diabetes	Clinicaltrials.gov NCT00541775	Merck	10/05/07	12/17/10	06/2006	03/2007	03/2007	05/17/10	03/2007

	Clinicaltrials.gov registration (or other registry) information									
Author, year	**Subgroups specified in the methods section**	**Registry Registry number**	**Study sponsor (as noted in registry)**	**First received date**	**Last updated**	**Study start date**	**Study completion date**	**Primary completion date (under tracking in clinicaltrials.gov)**	**Results first received**	**Primary completion date for recruitment**
Seino 2010	Previously treated with OAD therapy	Clinicaltrials.gov NCT00393718	Novo Nordisk	10/27/06	03/29/10	11/2006	05/2008	11/2007	02/23/10	05/2008
van der Meer 2009	None	Controlled-trials.com ISRCTN53177482	VU University Medical Centre Netherlands	12/20/05	05/11/10	09/01/04	09/01/06 Please note that the anticipated end date of this trial has been extended to 01/15/07.	NA (not in clinicaltrials.gov)	No study results posted	NA (not in clinicaltrials.gov)

Table B8. Oral Hypoglycemic Agents: Data Abstraction: Outcomes

	Clinicaltrials.gov registration (or other registry) information								
Author, year	Proposed/ Target N	Relevant original primary outcome in the registry	Relevant current primary outcome in the registry	Date of change in the relevant primary outcome	Original relevant secondary outcomes in the registry	Current relevant secondary outcomes in the registry	Date of change in the relevant secondary outcome	F/U 1 F/U 2 F/U 3 (weeks)	Results reported in the registry?
Aschner 2010	1050	HbA1c after 24 weeks	Change from baseline in hemoglobin A1c (HbA1c) at week 24	NA	None	None	NA	24 NA NA	Yes
Bakris 2006	336	Percent change from baseline in ACR after 32 weeks of treatment	Percent change from baseline in ACR after 32 weeks of treatment	NA	None	None	NA	32 NA NA	No
Bunck 2009	69	None	None	NA	None	Change in HbA1c from week 0 to week 52.	12/24/10	52 56 (reported for some secondary outcomes) NA	Yes
Defronzo 2009	1462	Change from baseline in HbA1c to week 24	Baseline and change from baseline in Hemoglobin A1c (A1C) at week 24	NA	None	Percentage of participants achieving therapeutic glycemic response (A1C < 7.0%) at week 24	07/15/11	24 NA NA	Yes
Defronzo 2010	137	None	None	NA	None	Change in HbA1c at week 20	07/21/09	20 NA NA	Yes

	Clinicaltrials.gov registration (or other registry) information								
Author, year	**Proposed/ Target N**	**Relevant original primary outcome in the registry**	**Relevant current primary outcome in the registry**	**Date of change in the relevant primary outcome**	**Original relevant secondary outcomes in the registry**	**Current relevant secondary outcomes in the registry**	**Date of change in the relevant secondary outcome**	**F/U 1 F/U 2 F/U 3 (weeks)**	**Results reported in the registry?**
Garber 2009	746	Change in A1c at 52 weeks	Change in A1c at week 52, 104, 156	NA#	Glycemic control	None	03/24/11	52 104 156	Yes
Goldberg 2005	719	None	None	NA	None	Change in A1c	12/11/08	"Anticipated to be about 39 weeks" NA NA	No
Goldstein 2007; Williams-Herman 2009	1208	HbA1c	Change from baseline in HbA1c (Hemoglobin A1C) at week 24	04/07/10	None	Change from baseline in HbA1c (Hemoglobin A1C) at week 54, week 104	NA#	24 54 104	Yes
Gupta 2009	60	None	None	NA	None	None	NA	4 months NA NA	No
Hamann 2008	544	Change in HbA1c level from baseline following 52 weeks of treatment	Change in HbA1c level from baseline following 52 weeks of treatment	NA	None	None	NA	52 NA NA	No

	Clinicaltrials.gov registration (or other registry) information								
Author, year	**Proposed/ Target N**	**Relevant original primary outcome in the registry**	**Relevant current primary outcome in the registry**	**Date of change in the relevant primary outcome**	**Original relevant secondary outcomes in the registry**	**Current relevant secondary outcomes in the registry**	**Date of change in the relevant secondary outcome**	**F/U 1 F/U 2 F/U 3 (weeks)**	**Results reported in the registry?**
Jadzinsky 2009	1306	Mean reduction in baseline A1C values after 24 weeks of treatment	Change from baseline in hemoglobin A1c (A1C) at week 24	NA	Subjects achieving a glycemic response defined as A1C < 7.0%	Percentage of participants achieving A1C < 7% and ≤6.5% at Week 24	06/30/10	24 NA NA	Yes
Kaku 2009	160	NR*	Change in HbA1C at the end of the treatment period	NA	None	None	NA	28 NA NA	Says yes but unable to locate
Nauck 2007; Seck 2010	1172	After 52 weeks, reduction in HbA1C	Change From Baseline in HbA1c at week 52	NA	Reduction in A1c at 104 weeks, durability of glycemic efficacy	Change from baseline in HbA1c at week 104	NA	52 104 NA	Yes
Nauck 2009	1091	HbA1c after 26 weeks of treatment	Change in Glycosylated A1c (HbA1c) at week 26 and week 104	NA#	None	None	NA	26 104 NA	Yes
Perez 2009	600	The change from baseline in hemoglobin Alc. [Time Frame: 24 Weeks]	Percent Change From baseline in Glycosylated Hemoglobin [Time Frame: Baseline and Week 24]	NA	None	None	NA	24 NA NA	Yes

	Clinicaltrials.gov registration (or other registry) information								
Author, year	**Proposed/ Target N**	**Relevant original primary outcome in the registry**	**Relevant current primary outcome in the registry**	**Date of change in the relevant primary outcome**	**Original relevant secondary outcomes in the registry**	**Current relevant secondary outcomes in the registry**	**Date of change in the relevant secondary outcome**	**F/U 1 F/U 2 F/U 3 (weeks)**	**Results reported in the registry?**
Pratley 2010	665	HbA1c (Time Frame: after 26 weeks of treatment)	Mean Change From Baseline in HbA1c at Week 26, 52, 78	NA#	None	Percentage of subjects achieving treatment target of HbA1c < 7.0%, < 6.5% at Week 26, 52, 78. Based on the FAS.	06/01/11	26 52 78	Yes
Raskin 2009	560	HbA1c after 26 weeks of treatment	HbA1c; Time Frame: after 26 weeks of treatment	NA	Percentage of subjects achieving sudden levels of HbA1c	None	NA	26 NA NA	No
Raz 2008	190	HbA1c after 18 weeks, safety and tolerability	Change from baseline in A1C at week 18	NA	None	Change in A1c at week 30	04/29/10	18 30 NA	Yes
Rigby 2009	169	Change in HbA1c from baseline to week 16 endpoint	Mean percentage of change in HbA1c from week 0 (Baseline) to week 16 endpoint	NA	Mean Percentage of Change in HbA1c	Mean percentage of change in HbA1c	NA	16 8 NA	Yes
Robbins 2007	320	HbA1C	HbA1C	NA	None	None	NA	NR NA NA	No

	Clinicaltrials.gov registration (or other registry) information								
Author, year	**Proposed/ Target N**	**Relevant original primary outcome in the registry**	**Relevant current primary outcome in the registry**	**Date of change in the relevant primary outcome**	**Original relevant secondary outcomes in the registry**	**Current relevant secondary outcomes in the registry**	**Date of change in the relevant secondary outcome**	**F/U 1 F/U 2 F/U 3 (weeks)**	**Results reported in the registry?**
Rosenstock 2006	453	Change from baseline in hemoglobin A1c (HbA1c) at week 32	Change from baseline in hemoglobin A1c (HbA1c) at week 32.	NA	HbA1c	HbA1c	NA	32 NA NA	No
Scott 2008	273	Hemoglobin A1C (A1C) at week 18	Hemoglobin A1C (A1C) at week 18	NA	None	None	NA	18 NA NA	Yes
Seino 2010	400	HbA1C after 24 weeks of treatment	HbA1c after 24 weeks of treatment	NA	A1c at 52 weeks	A1c at 52 weeks	NA	24 52 NA	Yes
van der Meer 2009	90	NR*	None	NA	None	None	NA	24 NA NA	No

Abbreviations: ADA, American Diabetes Association; ANCOVA, analysis of covariance; ANOVA, analysis of variance; APT, all patients treated; AUC, area under curve; BMI, body mass index; CI, confidence interval; FAS, full analysis set; ITT, intent to treat; LOCF, last observation carried forward; LSM, least squares mean; NA, not applicable; NI, non-inferiority; NR, not reported; RCT, randomized controlled trial; SD, standard deviation.

* Not reported because registry did not capture original outcome.

Open-label extension not considered a change in outcome.

Table B9. Lipid Modifying Agents: Data Abstraction: Study Funder and Conflicts of Interest

		Study characteristics, from the publication						
Author, year	**Journal**	**Study funder**	**Percentage of authors employed by pharmaceutical industry**	**Percentage of authors with COI from industry**	**What company makes intervention drug?**	**What company makes comparator drug?**	**Was there any assistance authoring the publication?**	**Did the publication indicate that the trial was registered?**
Ballantyne, 2005	American Heart Journal	Merck & Co./Schering Plough Pharmaceuticals	3/5 (60%)	2/2 (0%)	Ezetimibe/ simvastatin: Merck (Vytorin)	Atorvastatin: Pfizer (Lipitor)	NR	No
Ballantyne, 2007	The American Journal of Cardiology	NR; (clearly AstraZeneca)	2/7 (29%)	1/5 (20%)	Rosuvastatin: AstraZeneca (Crestor)	Combination Rosuvastatin/ Ezetimibe. Ezetimibe: Merck	Yes	No
Blagden, 2007	Current Medical Research and Opinion	Schering-Plough UK Limited	1/2 (50%)	0/1 (0%)	Ezetimibe: Merck atorvastatin: Pfizer (Lipitor) and generic	Placebo/ atorvastatin: generic	Yes	No
Catapano, 2006	Current Medical Research and Opinion	Merck & Co./Schering-Plough Pharmaceuticals	4/7 (57%)	0/3 (0%)	Ezetimibe/ simvastatin: Merck	Rosuvastatin: AstraZeneca (Crestor)	Yes	Yes
Conard, 2008	The American Journal of Cardiology	Merck & Co./Schering-Plough Pharmaceuticals	5/8 (63%)	3/3 (100%)	Ezetimibe: Merck plus Atorvastatin: Pfizer and generic	Atorvastatin: Pfizer and generic	Yes	No
Constance, 2007	Diabetes, Obesity and Metabolism	Merck & Co./Schering Plough Pharmaceuticals	5/8 (63%)	0/3 (0%)	Ezetimibe/ simvastatin: Merck (Vytorin)	Atorvastatin: Pfizer (Lipitor)	NR	No
Goldberg, 2006	Mayo Clinic Proceedings	Merck & Co./Schering Plough Pharmaceuticals	4/8 (50%)	4/4 (100%)	Ezetimibe/ simvastatin: Merck (Vytorin)	Atorvastatin: Pfizer (Lipitor)	Yes	No
Gouni-Berthold, 2008	Atherosclerosis	MSD Sharp & Dohme, Germany, and the Wilhelm-Doerenkamp Foundation, Cologne	0/13 (0%)	0/13 (0%)	Ezetimibe: Merck Simvastatin: Merck and generic	Combination ezetimibe/ simvastatin: Merck	NR	No

		Study characteristics, from the publication						
Author, year	**Journal**	**Study funder**	**Percentage of authors employed by pharmaceutical industry**	**Percentage of authors with COI from industry**	**What company makes intervention drug?**	**What company makes comparator drug?**	**Was there any assistance authoring the publication?**	**Did the publication indicate that the trial was registered?**
Landray, 2006	American Journal of Kidney Diseases	Merck & Co.	0	0	Ezetimibe: Merck	Simvastatin: Merck	NR	Yes
Leiter, 2008	The American Journal of Cardiology	Merck & Co./Schering-Plough Pharmaceuticals	5/8 (63%)	3/3 (100%)	Ezetimibe: Merck plus Atorvastatin: Pfizer and generic	Atorvastatin: Pfizer and generic	Yes	Yes
Patel, 2006	International Journal of Clinical Practice	Schering- Plough Pharmaceuticals	0/2 (0%)	1/2 (50%)	Ezetimibe/ simvastatin: Merck	Placebo/ simvastatin: generic	NR	No
Pearson, 2005	Mayo Clinic Proceedings	Merck & Co./Schering Plough Pharmaceuticals	3/6 (50%)	2/3 (66%)	Ezetimibe plus statin therapy: Merck (ezetimibe)	Placebo plus statin therapy: NA	No	No
Reckless, 2008	International Journal of Clinical Practice	Merck & Co./Schering Plough Pharmaceuticals	6/10 (60%)	Can't be determined	Ezetimibe/ simvastatin: Merck (Vytorin)	Statin drug: not specified	Yes	No
Roeters van Lennep, 2008	Current Medical Research and Opinion	Merck & Co.; Sharp and Dohme and Schering Plough Pharmaceuticals	0/6 (0%)	3/6 (50%)	Atorvastatin: Pfizer and generic Simvastatin: Merck	Ezetimibe/ simvastatin: Merck Ezetimibe monotherapy: Merck	NR	Yes

Table B10. Lipid Modifying Agents: Data Abstraction: Study Characteristics

	Publication, methods section									
Author, year	Study design	Intervention	Comparison	Total N	Primary outcome stated in the study (relevant to index outcome)	Relevant secondary outcomes (relevant to index outcome)	Followup intervals (weeks) F/U 1 F/U 2 F/U 3	Analysis set (definition from study) Definition of analysis set	How handled missing values	Subgroups specified in the methods section
Ballantyne, 2005	Parallel group trial	Ezetimibe/simvastatin	Atorvastatin	1902	None	None	6 NA NA	MITT All randomized patients who had a valid baseline and at least one valid post baseline measurement.	NR	None
Ballantyne, 2007	Parallel group RCT	Rosuvastatin	Rosuvastatin and ezetimibe	469	None	None	6 NA NA	ITT All patients with a baseline lipid measurement and one lipid measurement after baseline and who had used one dose of study medication.	Last-available-observation-carried-forward	None
Blagden, 2007	Parallel group RCT	Ezetimibe plus atorvastatin	Placebo/atorvastatin	148	None	None	6 NA NA	ITT All randomized subjects who received at least one dose of study medication, and had at least one post-baseline measurement.	NR	None

	Publication, methods section									
Author, year	**Study design**	**Intervention**	**Comparison**	**Total N**	**Primary outcome stated in the study (relevant to index outcome)**	**Relevant secondary outcomes (relevant to index outcome)**	**Followup intervals (weeks)** **F/U 1** **F/U 2** **F/U 3**	**Analysis set (definition from study)** **Definition of analysis set**	**How handled missing values**	**Subgroups specified in the methods section**
Catapano, 2006	Parallel group RCT	Ezetimibe/simvastatin	Rosuvastatin	2959	None	None	6 NA NA	Efficacy outcome: MITT population was used. Safety outcome: all patients who received at least one dose of the double-blind study medication. Efficacy outcome: all randomized patients who had a valid baseline and at least one valid post-baseline measurement.	NR	NR

	Publication, methods section									
Author, year	**Study design**	**Intervention**	**Comparison**	**Total N**	**Primary outcome stated in the study (relevant to index outcome)**	**Relevant secondary outcomes (relevant to index outcome)**	**Followup intervals (weeks)** **F/U 1** **F/U 2** **F/U 3**	**Analysis set (definition from study)** **Definition of analysis set**	**How handled missing values**	**Subgroups specified in the methods section**
Conard, 2008	Parallel group RCT	Ezetimibe plus atorvastatin	Atorvastatin	196	None	None	6 NA NA	Full analysis-set approach for efficacy outcome. Safety outcome assessed in all treated patients with >=1 on-treatment measurement. All randomized patients who took 1 dose of study medication and had baseline and 1 post baseline values. Post baseline measurements up to 3 days after the last dose of double-blind study medication were included in the analysis.	NR	Age, sex, BMI, region, Bl LDL, etc.

	Publication, methods section									
Author, year	**Study design**	**Intervention**	**Comparison**	**Total N**	**Primary outcome stated in the study (relevant to index outcome)**	**Relevant secondary outcomes (relevant to index outcome)**	**Followup intervals (weeks)** **F/U 1** **F/U 2** **F/U 3**	**Analysis set (definition from study)** **Definition of analysis set**	**How handled missing values**	**Subgroups specified in the methods section**
Constance, 2007	Parallel group RCT	Ezetimibe/simvastatin	Atorvastatin	661	None	None	6 NA NA	All patients–treated approach for efficacy outcome. For safety: randomized patients who received at least 1 dose of study medication. Those patients who received at least one dose of randomized treatment, had a lipid measurement at baseline and had at least one lipid measurement following the start of treatment.	NR	Numerous; found on page 578
Goldberg, 2006	Parallel group trial	Ezetimibe/simvastatin	Atorvastatin	1229	None	None	6 NA NA	MITT for safety outcome: all randomized with 1+ dose of study medication. For efficacy: all randomized patients who had valid baseline measurements and at least one valid post baseline measurement.	NR	NR

	Publication, methods section									
Author, year	**Study design**	**Intervention**	**Comparison**	**Total N**	**Primary outcome stated in the study (relevant to index outcome)**	**Relevant secondary outcomes (relevant to index outcome)**	**Followup intervals (weeks)** **F/U 1** **F/U 2** **F/U 3**	**Analysis set (definition from study)** **Definition of analysis set**	**How handled missing values**	**Subgroups specified in the methods section**
Gouni-Berthold, 2008	Parallel group RCT	Ezetimibe versus simvastatin	Ezetimibe plus simvastatin	72	None	None	2 NA NA	NR NR	NR	None
Landray, 2006	Parallel group	Simvastatin plus ezetimibe	Simvastatin plus placebo ezetimibe	203	None	All cause mortality (according to CER), but this outcome is not reported in the publication. "Serious adverse events"	4 6	ITT All patients allocated to simvastatin plus ezetimibe versus all those allocated to simvastatin plus placebo ezetimibe. For safety: NR, but all patients analyzed.	Missing followup blood results were inputted with the value recorded at randomization.	None

	Publication, methods section									
Author, year	**Study design**	**Intervention**	**Comparison**	**Total N**	**Primary outcome stated in the study (relevant to index outcome)**	**Relevant secondary outcomes (relevant to index outcome)**	**Followup intervals (weeks)** **F/U 1** **F/U 2** **F/U 3**	**Analysis set (definition from study)** **Definition of analysis set**	**How handled missing values**	**Subgroups specified in the methods section**
Leiter, 2008	Parallel group RCT	Ezetimibe plus atorvastatin	Atorvastatin	579	None	None	6 NA NA	FAS Efficacy outcome: all randomly assigned patients who used 1 dose of study medication and had a baseline value and 1 post baseline value. Safety outcome: all patients randomly assigned and used >=1 dose of study medication were included in the safety analyses.	NR	Numerous. Example: age, sex, BMI, baseline lipids, etc.
Patel, 2006	Parallel group trial	Ezetimibe and simvastatin	Simvastatin plus placebo	153	None	None	6 NA NA	Safety population Efficacy; ITT and per-protocol The safety population included all subjects who were randomized and received at least one dose of study medication, and was used for the safety analysis.	NR	NR

	Publication, methods section									
Author, year	**Study design**	**Intervention**	**Comparison**	**Total N**	**Primary outcome stated in the study (relevant to index outcome)**	**Relevant secondary outcomes (relevant to index outcome)**	**Followup intervals (weeks)** **F/U 1** **F/U 2** **F/U 3**	**Analysis set (definition from study)** **Definition of analysis set**	**How handled missing values**	**Subgroups specified in the methods section**
Pearson, 2005	Parallel group trial	Ezetimibe plus their current statin therapy and dose	Placebo, plus their current statin therapy and dose	3030	None	None	6 NA NA	MITT for efficacy Safety: all patients who received the study drug. All randomized patients with a baseline assessment and at least 1 valid post baseline assessment of LDL-C level was used for the effectiveness analyses. All patients who received the study drug were included in the safety analyses.	NR	3 NCEP ATP II risk categories

	Publication, methods section									
Author, year	**Study design**	**Intervention**	**Comparison**	**Total N**	**Primary outcome stated in the study (relevant to index outcome)**	**Relevant secondary outcomes (relevant to index outcome)**	**Followup intervals (weeks)** **F/U 1** **F/U 2** **F/U 3**	**Analysis set (definition from study)** **Definition of analysis set**	**How handled missing values**	**Subgroups specified in the methods section**
Reckless, 2008	Parallel group RCT	Ezetimibe／simvastatin	Fluvastatin; lovastatin; pravastatin; simvastatin; atorvastatin; rosuvastatin	424	None	None	12 NA NA	FAS for efficacy outcome; all-patients-as-treated population for safety. Patients who took at least one dose of randomized treatment, had a lipid measurement at baseline and at least one lipid measurement following the start of treatment. Safety population; all as treated: all randomized patients who took at least one dose of the open-label study medication.	Last value carried forward	None

	Publication, methods section									
Author, year	**Study design**	**Intervention**	**Comparison**	**Total N**	**Primary outcome stated in the study (relevant to index outcome)**	**Relevant secondary outcomes (relevant to index outcome)**	**Followup intervals (weeks)** **F/U 1** **F/U 2** **F/U 3**	**Analysis set (definition from study)** **Definition of analysis set**	**How handled missing values**	**Subgroups specified in the methods section**
Roeters van Lennep, 2008	Parallel group RCT	Doubling statin dose (either simvastatin or atorvastatin)	Ezetimibe/ simvastatin combination tablet	367	None	"Safety and tolerability ... also assessed throughout the study", including fatal events	12 NA NA	ITT for efficacy analysis; no discussion of population for safety analyses. Efficacy: all randomized patients. Safety; none defined, but all randomized patients are presented in the adverse event table.	NR	NR

Table B11. Lipid Modifying Agents: Data Abstraction: Study Registration

	Clinicaltrials.gov registration (or other registry) information								
Author, year	Registry Registry number	Study sponsor (as noted in registry)	First received date	Last updated	Study start date	Study completion date	Primary completion date (under tracking in clinicaltrials.gov)	Results first received	Primary completion date for recruitment
Ballantyne, 2005	ClinicalTrials.gov NCT00092690	Merck	09/23/04	01/21/10	01/2003	NR	NR	No study results posted	Not stated
Ballantyne, 2007	ClinicalTrials.gov NCT00653445	AstraZeneca	04/02/08	03/25/09	06/2004	06/2005	NR	No study results posted	Not stated
Blagden, 2007	ISRCTN ISRCTN18808154	Schering-Plough UK Ltd	08/13/04	06/21/11	01/01/04	12/31/04	NR	No study results posted	Not an option in ISCRTN registry
Catapano, 2006	ClinicalTrials.gov NCT00090298	Merck	08/25/04	01/21/10	05/2004	NR	NR	No study results posted	NR
Conard, 2008	ClinicalTrials.gov NCT00276458	Merck	01/10/06	04/14/10	02/2006	02/2008	01/2008	12/18/08	01/08
Constance, 2007	ClinicalTrials.gov NCT00093106 duplicate with: NCT00541697	Merck	10/05/07	09/23/09	02/2005	10/2005	NR	No study results posted	Not stated
Goldberg, 2006	ClinicalTrials.gov NCT00110435	Merck	05/09/05	01/27/10	06/2005	04/2006	04/2006	No study results posted	04/2006
Gouni-Berthold, 2008	ClinicalTrials.gov NCT00317993	University of Cologne	04/24/06	NA ("no posted changes")	04/2004	07/2004	NR	No study results posted	Not stated
Landray, 2006	Cochrane Renal Group Registry of Clinical Trials CRG060500006	NR	NR	NR	02/2002	02/2003	NR	NR	NA

	Clinicaltrials.gov registration (or other registry) information								
Author, year	**Registry** **Registry number**	**Study sponsor (as noted in registry)**	**First received date**	**Last updated**	**Study start date**	**Study completion date**	**Primary completion date (under tracking in clinicaltrials.gov)**	**Results first received**	**Primary completion date for recruitment**
Leiter, 2008	ClinicalTrials.gov NCT00276484	Merck	01/11/06	04/14/10	02/2006	03/2008	03/2008	02/06/09	03/08
Patel, 2006	ISRCTN ISRCTN47214063	Schering-Plough UK Ltd	08/13/04	11/17/10	04/01/00	05/21/05	NR	No study results posted	Not an option in ISCRTN registry
Pearson, 2005	ClinicalTrials.gov NCT00092586	Merck	09/23/04	01/21/10	09/2002	NR	NR	No study results posted	Not stated
Reckless, 2008	ClinicalTrials.gov NCT00132717	Merck	08/02/05	01/21/10	01/2005	NR	NR	No study results posted	Not stated
Roeters van Lennep, 2008	ClinicalTrials.gov NCT00166530	Merck	09/09/05	09/05/08	11/2005	NR	02/2007	No study results posted	Not stated

Table B12. Lipid Modifying Agents: Data Abstraction: Outcomes

	Clinicaltrials.gov registration (or other registry) information								
Author, year	Proposed/ Target N	Relevant original primary outcome in the registry	Relevant current outcome in the registry	Date of change in the primary outcome	Original relevant secondary outcomes in the registry	Current relevant secondary outcomes in the registry	Date of change in the relevant secondary outcome	F/U 1 F/U 2 F/U 3 (weeks)	Results reported in the registry?
Ballantyne, 2005	1640	None	None	NA	None	None	NA	6 NA NA	No
Ballantyne, 2007	NR	None	None	NA	None	None	NA	6 NA NA	No
Blagden, 2007	"Not provided at time of registration"	None	None	NA	None	None	NA	6 NA NA	No
Catapano, 2006	2725	None	None	NA	None	None	NA	6 NA NA	No
Conard, 2008	196	None	None	NA	None	None	NA	6 NA NA	Yes
Constance, 2007	500	None	None	NA	None	None	NA	6 NA NA	No

	Clinicaltrials.gov registration (or other registry) information								
Author, year	Proposed/ Target N	Relevant original primary outcome in the registry	Relevant current outcome in the registry	Date of change in the primary outcome	Original relevant secondary outcomes in the registry	Current relevant secondary outcomes in the registry	Date of change in the relevant secondary outcome	F/U 1 F/U 2 F/U 3 (weeks)	Results reported in the registry?
Goldberg, 2006	1125	None	None	NA	None	None	NA	6 NA NA	No
Gouni-Berthold, 2008	60	None	None	NA	None	None	NA	2 NA NA	No
Landray, 2006	NR	None	None	NA	None	None	NA	4 12 24	No
Leiter, 2008	579	None	None	NA	None	None	NA	6 NA NA	Yes; death NR
Patel, 2006	"Not provided at time of registration"	None	None	NA	None	None	NA	6 NA NA	No
Pearson, 2005	3000	None	None	NA	None	None	NA	6 NA NA	No

	Clinicaltrials.gov registration (or other registry) information								
Author, year	**Proposed/ Target N**	**Relevant original primary outcome in the registry**	**Relevant current outcome in the registry**	**Date of change in the primary outcome**	**Original relevant secondary outcomes in the registry**	**Current relevant secondary outcomes in the registry**	**Date of change in the relevant secondary outcome**	**F/U 1 F/U 2 F/U 3 (weeks)**	**Results reported in the registry?**
Reckless, 2008	450	None	None	NA	None	None	NA	12 NA NA	No
Roeters van Lennep, 2008	367	None	None	NA	None	None	NA	12 NA NA	No

Abbreviations: ANCOVA, analysis of covariance; ANOVA, analysis of variance; BMI, body mass index; FAS, full analysis set; F/U, follow up; ITT, intent to treat; LSM, least squares mean; MITT, modified intent to treat; NA, Not applicable; NR, Not reported; RCT, Randomized controlled trial; SD, standard deviation; TC, total cholesterol.

* Not reported because registry did not capture original outcome.

Open-label extension not considered a change in outcome.

Table B13. Osteoporosis: Data Abstraction: Study Funder and Conflicts of Interest

		Study Characteristics						
Author, year	**Journal**	**Study funder**	**Percentage of authors employed by pharmaceutical industry**	**Percentage of authors with COI from industry**	**What company makes intervention drug?**	**What company makes comparator drug?**	**Was there any assistance authoring the publication?**	**Did the publication indicate that the trial was registered?**
Barrett-Connor, 2006	New England Journal of Medicine	Eli Lilly	2/8 (25%)	6/6 (100%)	Raloxifene: Eli Lilly	Placebo: NA	Yes	Yes
Black, 2007	New England Journal of Medicine	Novartis Pharmaceuticals	7/21 (33%)	13/14 (93%)	Zoledronic Acid: Novartis Pharmaceuticals	Placebo: NA	No	Yes
Bonnick, 2006 NOTE: this is a companion to Rosen 2005	The Journal of Clinical Endocrinology & Metabolism	Merck	3/11 (27%)	7/8 (88%)	Alendronate: Merck and generic	Risedronate: Warner Chilcott	Yes	Yes
Grant, 2005	Lancet	The UK Medical Research Council funded the central organization of RECORD	0/14 (0%)	3/14 (21%)	Vitamin D3: NA calcium: NA (Shire Pharmaceutical funded the drugs, manufactured by Nycomed)	Placebo: NA	No	NR

		Study Characteristics						
Author, year	**Journal**	**Study funder**	**Percentage of authors employed by pharmaceutical industry**	**Percentage of authors with COI from industry**	**What company makes intervention drug?**	**What company makes comparator drug?**	**Was there any assistance authoring the publication?**	**Did the publication indicate that the trial was registered?**
Greenspan, 2006	The Journal of Clinical Endocrinology & Metabolism	National Institutes of Health/National Institute of Diabetes and Digestive and Kidney Diseases (K24 DK062895-03), a NCST from Procter and Gamble and the Alliance for Better Bone Health and to the General Clinical Research Center of the University of Pittsburgh by the National Institutes of Health/National Center for Research Resources (M01-RR00056)	0/5 (0%)	4/5 (80%)	Risedronate: Warner Chilcott Article says it was provided by Procter and Gamble	Placebo: NA Calcium and Vitamin D supplements: Provided by GlaxoSmithKline Article says it was provided by Procter and Gamble	NR	No
Jackson, 2006	New England Journal of Medicine	Supported by the National Heart, Lung, and Blood Institute and the General Clinical Research Center program of the National Center for Research Resources, Department of Health and Human Services. The active study drug and placebo were supplied by Glaxo SmithKline Consumer Healthcare (Pittsburgh).	0/47 (0%)	10/47 (21%)	Calcium carbonate and vitamin D3: generic The active study drug and placebo were supplied by GlaxoSmithKline Consumer Healthcare	Placebo: NA The active study drug and placebo were supplied by GlaxoSmithKline Consumer Healthcare	No	Yes

		Study Characteristics						
Author, year	**Journal**	**Study funder**	**Percentage of authors employed by pharmaceutical industry**	**Percentage of authors with COI from industry**	**What company makes intervention drug?**	**What company makes comparator drug?**	**Was there any assistance authoring the publication?**	**Did the publication indicate that the trial was registered?**
McClung, 2006	New England Journal of Medicine	Amgen	4/16 (25%)	11/12 (92%)	Denosumab: Amgen	Alendronate: generic, Merck and placebo	Yes	Yes
Porthouse, 2005	British Medical Journal	Grants from Northern and Yorkshire NHS research and development, healthy ageing programme (TA, RMF, AS, IW, DJT), Shire, and Nycomed.	0/15 (0%)	5/15 (33%)	Combination calcium and cholecalciferol (vitamin D3): both supplied by Shire	NA	NR	Yes
Prince, 2006	Archives of Internal Medicine	Healthway Health Promotion Foundation of Western Australia and by project grant 254627 from the National Health and Medical Research Council of Australia.	0/4 (0%)	0/4 (0%)	Calcium carbonate tablets: multiple pharmaceutical companies and generic	Placebo tablets: NA	No	No
Reid, 2006	The American Journal of Medicine	NR	0/8 (0%)	0/8 (0%)	Calcium citrate: generic Provided by Mission Pharmacal	Placebo: NA Provided by Mission Pharmacal	Yes	Yes
Rosen, 2005 NOTE: this is a companion to Bonnick 2006 but has separate NCT number	Journal of Bone and Mineral Research	NR (clearly Merck)	4/11 (36%)	7/7 (100%)	Alendronate: Merck and generic	Risedronate: Warner Chilcott	No	No

		Study Characteristics						
Author, year	**Journal**	**Study funder**	**Percentage of authors employed by pharmaceutical industry**	**Percentage of authors with COI from industry**	**What company makes intervention drug?**	**What company makes comparator drug?**	**Was there any assistance authoring the publication?**	**Did the publication indicate that the trial was registered?**
Vogel, 2006	Journal of the American Medical Association	Public Health Service grants from the National Cancer Institute, National Institutes of Health, Department of Health and Human Services; and by AstraZeneca Pharmaceuticals and Eli Lilly and Co.	0/21 (0%)	5/21 (24%)	Tamoxifen: Cytogen and generic	Raloxifene: Eli Lilly	Yes	Yes

Table B14. Osteoporosis: Data Abstraction: Study Characteristics

	Publication, methods section									
Author, year	**Study design**	**Intervention**	**Comparison**	**Total N**	**Primary outcome as stated in the study (relevant to index outcome)**	**Relevant secondary outcomes (relevant to index outcome)**	**Followup intervals (months or years)** **F/U 1** **F/U 2** **F/U 3**	**Analysis set (definition from study)** **Definition of analysis set**	**How handled missing values**	**Subgroups specified in the methods section**
Barrett-Connor, 2006	Parallel group	Raloxifene	Placebo	10,101	None	Fracture; clinical nonvertebral and vertebral	Median 5.6 years range, 0.01 to 7.06 NA NA	ITT NR	NA (time-to-event data for primary outcomes)	None for fractures
Black, 2007	Parallel group placebo RCT	Zoledronic acid	Placebo	3889	New vertebral fracture (in patients not taking concomitant osteoporosis medications) and hip fracture (in all patients)	Secondary efficacy endpoints: any nonvertebral fracture, any clinical fracture, and clinical vertebral fracture	12 m 24 m 36 m	Efficacy analyses included all patients who had undergone randomization except for 29 whose site was terminated. The incidence of vertebral fracture included patients who had undergone radiography at baseline at least once during F/U.	NR	NR

	Publication, methods section									
Author, year	**Study design**	**Intervention**	**Comparison**	**Total N**	**Primary outcome as stated in the study (relevant to index outcome)**	**Relevant secondary outcomes (relevant to index outcome)**	**Followup intervals (months or years)** **F/U 1** **F/U 2** **F/U 3**	**Analysis set (definition from study)** **Definition of analysis set**	**How handled missing values**	**Subgroups specified in the methods section**
Bonnick, 2006 NOTE: this is a companion to Rosen 2005	Parallel group RCT, extension study of Rosen 2005	Alendronate	Risedronate	833	None	None	12 month extension after initial 12 months NA NA	For safety outcomes, all patients who received at least one dose of study medication in the extension period	NA for safety outcomes	None
Grant, 2005	Factorial design, parallel group	Oral vitamin D3 combined with calcium	Placebo	5292	All-new low-energy fractures including clinical, radiologically confirmed vertebral fractures, but not those of the face or skull	None	24 to 64 m NA NA	ITT NR	NR	High or low weight (less than 55 kg or not); latitude of recruitment center; dietary calcium; and vitamin D exposure from the sun or diet
Greenspan, 2006	Parallel group	Risedronate	Placebo	87	None	None	12 m 24 m (extension) NA	ITT NR	NR	None

	Publication, methods section									
Author, year	**Study design**	**Intervention**	**Comparison**	**Total N**	**Primary outcome as stated in the study (relevant to index outcome)**	**Relevant secondary outcomes (relevant to index outcome)**	**Followup intervals (months or years)** **F/U 1** **F/U 2** **F/U 3**	**Analysis set (definition from study)** **Definition of analysis set**	**How handled missing values**	**Subgroups specified in the methods section**
Jackson, 2006	Parallel group	Elemental calcium as calcium carbonate with vitamin D3	Placebo	36,282	Total fractures defined as all reported clinical fractures other than ribs, face, etc.	None	7 years average NA NA	Time-to-event basis according to the ITT principle NR	NR	
McClung, 2006	Parallel group, placebo control and active control RCT	Denosumab; alendronate	Placebo	412	None	None	12 m NA NA	Efficacy analyses: ITT Fractures were reported as a safety outcome and that analysis set was no specified explicitly (was n=406 from adverse event table). All subjects with a baseline value and at least one value after baseline and compared across dose groups.	NR	NA

	Publication, methods section									
Author, year	**Study design**	**Intervention**	**Comparison**	**Total N**	**Primary outcome as stated in the study (relevant to index outcome)**	**Relevant secondary outcomes (relevant to index outcome)**	**Followup intervals (months or years)** **F/U 1** **F/U 2** **F/U 3**	**Analysis set (definition from study)** **Definition of analysis set**	**How handled missing values**	**Subgroups specified in the methods section**
Porthouse, 2005	Parallel group RCT, open label	Calcium with cholecalciferol and information leaflet on dietary calcium intake and prevention of falls	Leaflet only	3454	All clinical fractures	Hip fractures	25 months (range 18 to 42 months) NA NA	ITT NR	NR	Hip and wrist fractures
Prince, 2006	Parallel group	Calcium carbonate	Placebo	1460	Clinical incident osteoporotic fractures, vertebral deformity, and adverse events ascertained in 5 years	None	5 y NA NA	ITT NR	NR	Patients consuming 80% or more of tablets

	Publication, methods section									
Author, year	**Study design**	**Intervention**	**Comparison**	**Total N**	**Primary outcome as stated in the study (relevant to index outcome)**	**Relevant secondary outcomes (relevant to index outcome)**	**Followup intervals (months or years)** **F/U 1** **F/U 2** **F/U 3**	**Analysis set (definition from study)** **Definition of analysis set**	**How handled missing values**	**Subgroups specified in the methods section**
Reid, 2006	Parallel group	Calcium	Placebo	1471	Time to first clinical fracture at any site	Fracture subgroups: total vertebral fractures, hip fractures, distal forearm fractures, and osteoporotic fractures (comprising all fractures except those of the head, hands, feet, and ankles, and resulting from major trauma).	"Over 5 years" NA NA	ITT and per protocol Per protocol pre-specified as primary analysis "because of the likelihood that other anti-osteoporotic therapies would have much greater effects on bone density and fracture than calcium…" NR	NR	Total vertebral fractures, hip fractures, distal forearm fractures, and osteoporotic fractures
Rosen, 2005 NOTE: this is a companion to Bonnick 2006 but has separate NCT number	Parallel group RCT	Alendronate	Risedronate	1053	None	None	6 m 12 m NA	ITT All patients who received at least one dose of study drug in either treatment group for safety analyses	LOCF	None

	Publication, methods section									
Author, year	**Study design**	**Intervention**	**Comparison**	**Total N**	**Primary outcome as stated in the study (relevant to index outcome)**	**Relevant secondary outcomes (relevant to index outcome)**	**Followup intervals (months or years)** **F/U 1** **F/U 2** **F/U 3**	**Analysis set (definition from study)** **Definition of analysis set**	**How handled missing values**	**Subgroups specified in the methods section**
Vogel, 2006	Parallel group RCT	Tamoxifen	Raloxifene	19747	None	Osteoporotic fractures	5 y NA NA	ITT All randomized participants with followup data who were at risk at baseline for the diagnosis of an incident case of breast cancer	NR	None

Table B15. Osteoporosis: Data Abstraction: Study Registration

	Clinicaltrials.gov registration (or other registry) information								
Author, year	**Registry Registry number**	**Study sponsor (as noted in registry)**	**First received date**	**Last updated**	**Study start date**	**Study completion date**	**Primary completion date (under tracking in clinicaltrials.gov)**	**Results first received**	**Primary completion date for recruitment**
Barrett-Connor, 2006	clinicaltrials.gov NCT00190593	Eli Lilly	09/12/05	01/24/07	06/1998	11/2005	NR	No study results posted	Not stated
Black, 2007	clinicaltrials.gov NCT00049829	Novartis Pharmaceuticals	11/14/02	11/01/11	01/2002	NR	06/2006	No study results posted	06/2006
Bonnick, 2006 NOTE: this is a companion to Rosen 2005	clinicaltrials.gov NCT00092014	Merck	09/21/04	01/21/10	02/2002	NR	NR	No study results posted	Not stated
Grant, 2005	Not stated (clearly controlled-trials.com) ISRCTN51647438	Medical Research Council (MRC) (UK)	10/23/00	07/22/09	11/18/98	04/30/04	NR	No study results posted	Not stated
Greenspan, 2006	clinicaltrials.gov NCT00118508	University of Pittsburgh	06/30/05	08/09/11	05/2003	07/2006	07/2006	No study results posted	07/2006
Jackson, 2006	clinicaltrials.gov NCT00000611	National Heart, Lung, and Blood Institute (NHLBI)	10/27/99	11/27/06	NR	NR	NR	NR	Not stated
McClung, 2006	clinicaltrials.gov NCT00043186	Amgen	08/06/02	06/25/10	05/2002	06/2007	04/2007	12/22/09	04/2007

	Clinicaltrials.gov registration (or other registry) information								
Author, year	**Registry** **Registry number**	**Study sponsor (as noted in registry)**	**First received date**	**Last updated**	**Study start date**	**Study completion date**	**Primary completion date (under tracking in clinicaltrials.gov)**	**Results first received**	**Primary completion date for recruitment**
Porthouse, 2005	ISRCTN www.controlled-trials.com ISRCTN26118436	NHS R&D Regional Programme Register - Department of Health (UK)	01/23/04	06/07/11	01/09/01	01/03/04	NR	No study results posted	Not stated
Prince, 2006	Australian Clinical Trials Registry www.anzctr.org.au ACTRN12607000055404	Primary: Individual; Secondary: University of Western Australia and Sir Charles Gairdner Hospital	01/11/07 (submitted) 01/17/07 (registered)	"Trial not updated since registration"	21/07/1998	NR (must be around 01/2007)	NR	No study results posted	Not available in this registry
Reid, 2006	Australian Clinical Trials Registry www.anzctr.org.au ACTRN 012605000242628	Primary: Individual (Professor Reid); Secondary: University of Auckland Bone Research Group	08/24/05 (submitted) 08/31/205 (registered)	"Trial not updated since registration"	03/01/98	NR (must be around 08/2005)	NR	No study results posted	Not stated

	Clinicaltrials.gov registration (or other registry) information								
Author, year	**Registry** **Registry number**	**Study sponsor (as noted in registry)**	**First received date**	**Last updated**	**Study start date**	**Study completion date**	**Primary completion date (under tracking in clinicaltrials. gov)**	**Results first received**	**Primary completion date for recruitment**
Rosen, 2005 NOTE: this is a companion to Bonnick 2006 but has separate NCT number	clinicaltrials.gov NCT00092040	Merck	09/21/04	01/21/10	03/2003	NR	NR	No study results posted	Not stated
Vogel, 2006	clinicaltrials.gov NCT00003906	National Surgical Adjuvant Breast and Bowel Project (NSABP)	11/01/99	09/20/11	05/1999	03/2014	12/2005	No study results posted	03/2014

Table B16. Osteoporosis: Data Abstraction: Outcomes

	Clinicaltrials.gov registration (or other registry) information								
Author, year	**Proposed/ Target N**	**Relevant original primary outcome in the registry**	**Relevant current outcome in the registry**	**Date of change in the primary outcome**	**Original relevant secondary outcomes in the registry**	**Current relevant secondary outcomes in the registry**	**Date of change in the relevant secondary outcomes**	**F/U 1** **F/U 2** **F/U 3** **(months or years)**	**Results reported in the registry?**
Barrett-Connor, 2006	10000	None	None	NA	Fractures	Fractures	NA	5 to 7.5 y NA NA	No
Black, 2007	7700	Incidence of hip fractures Incidence of new vertebral fractures	Incidence of hip fractures Incidence of new vertebral fractures	NA	New and/or worsening vertebral fractures; all clinical fractures	New and/or worsening vertebral fractures; all clinical fractures	NA	NR NA NA	No
Bonnick, 2006 NOTE: this is a companion to Rosen 2005	900	None	None	NA	None	None	NA	24 m NA NA	No
Grant, 2005	5250	NR*	New fractures	NA	NR*	Not provided at time of registration	NA	NR NA NA	No
Greenspan, 2006	87	None	None	NA	None	None	NA	6 m 12 m 18 m; 24 m	No
Jackson, 2006	NR	None	None	NA	None	None	NA	NR NA NA	No

	Clinicaltrials.gov registration (or other registry) information								
Author, year	**Proposed/ Target N**	**Relevant original primary outcome in the registry**	**Relevant current outcome in the registry**	**Date of change in the primary outcome**	**Original relevant secondary outcomes in the registry**	**Current relevant secondary outcomes in the registry**	**Date of change in the relevant secondary outcomes**	**F/U 1** **F/U 2** **F/U 3** **(months or years)**	**Results reported in the registry?**
McClung, 2006	412	None	None	NA	None	None	NA	12 m 24 m; 36 m 42 m; 48 m	Yes
Porthouse, 2005	3314	NR*	All clinical fractures	NA	None	None	NA	NR NA NA	No
Prince, 2006	120	None	None	NA	None	None	NA	1 y 3 y 5 y	No
Reid, 2006	1500	Time to first clinical fracture	Time to first clinical fracture	NA	NR*	Total vertebral fractures, hip fractures, forearm fractures, osteoporotic fractures	NA	5 y NA NA	No
Rosen, 2005 NOTE: this is a companion to Bonnick 2006 but has separate NCT number	760	None	None	NA	None	None	NA	12 m NA NA	No

	Clinicaltrials.gov registration (or other registry) information								
Author, year	**Proposed/ Target N**	**Relevant original primary outcome in the registry**	**Relevant current outcome in the registry**	**Date of change in the primary outcome**	**Original relevant secondary outcomes in the registry**	**Current relevant secondary outcomes in the registry**	**Date of change in the relevant secondary outcomes**	**F/U 1 F/U 2 F/U 3 (months or years)**	**Results reported in the registry?**
Vogel, 2006	19747	None	None	NA	Effect of the therapy on the incidence of fractures of the hip, spine, or Colles' fractures of the wrist	Effect of the therapy on the incidence of fractures of the hip, spine, or Colles' fractures of the wrist	NA	5 y NA NA	No

Abbreviations: ANCOVA, analysis of covariance; ANOVA, analysis of variance; AUC, area under curve; BMD, bone mineral density; BMI, body mass index; CI, confidence interval; FAS, full analysis set populations; FPG, fasting plasma glucose; F/U, follow up; ITT, intent to treat; LOCF, last-observation-carried forward; MI, myocardial infarction; NA, not applicable; NI, non-inferiority; NR, not reported; RCT, randomized controlled trial; RR, risk ratio; SD, standard deviation.

* Not reported because registry did not capture original outcome.

Open-label extension not considered a change in outcome.

Figure B1. Oral Hypoglycemic Agents: Timelines of important dates in registered trials [source of date]

Aschner, 2010	**Date**	**Time** →	→	→	→
Study registered/first received [registry]	March 19,2007	✘			
Study start [registry]	March 2007	✘			
Recruitment started [publication]	NR				
Recruitment completed [publication]	NR				
Primary outcome completed [registry]	July 2008	✘			
Study completed [registry]	July 2008	✘			
Results first received [registry]	April 23,2009		✘		
Publication submitted [publication]	October 2, 2009			✘	
Publication accepted [publication]	November 19, 2009				✘
Article first published online [publication]	November 25,2009				✘
Publication printed [publication]	March 2010				✘
Registry last updated [registry]	April 20,2010				✘
Primary outcome changed [registry]	NA				

Bakris, 2006	**Date**	**Time** →	→	→
Study registered/first received [registry]	July 12,2007			✘
Study start [registry]	April 2000	✘		
Recruitment started [publication]	NR			
Recruitment completed [publication]	NR			
Primary outcome completed [registry]	June 2004	✘		
Study completed [registry]	June 2004	✘		
Results first received [registry]	None posted			
Publication submitted [publication]	February 8,2006		✘	
Publication accepted [publication]	May 17,2006			✘
Article first published online [publication]	NR			
Publication printed [publication]	October 2006			✘
Registry last updated [registry]	October 1,2010			✘
Primary outcome changed [registry]	NA			

Bunck, 2009	**Date**	**Time** → → → →
Study registered/first received [registry]	November 24,2004	✘
Study start [registry]	September 2004	✘
Recruitment started [publication]	September 27, 2004	✘
Recruitment completed [publication]	September 13,2007	✘
Primary outcome completed [registry]	December 2009	✘
Study completed [registry]	December 2009	✘
Results first received [registry]	December 24,2010	✘
Publication submitted [publication]	October 1,2008	✘
Publication accepted [publication]	January 19,2009	✘
Article first published online [publication]	February 5,2009	✘
Publication printed [publication]	May 2009	✘
Registry last updated [registry]	December 24,2010	✘
Primary outcome changed [registry]	December 24,2010	✘

Defronzo, 2009	**Date**	**Time** → → → → →
Study registered/first received [registry]	July 18,2005	✘
Study start [registry]	August 2005	✘
Recruitment started [publication]	NR	
Recruitment completed [publication]	NR	
Primary outcome completed [registry]	October 2006	✘
Study completed [registry]	February 2010	✘
Results first received [registry]	March 15,2011	✘
Publication submitted [publication]	November 3, 2008	✘
Publication accepted [publication]	May 21,2009	✘
Article first published online [publication]	May 28,2009	✘
Publication printed [publication]	September 2009	✘
Registry last updated [registry]	August 5,2011	✘
Primary outcome changed [registry]	NA	

Note- trials that reported more dates have wider displays.

Defronzo, 2010	**Date**	**Time** → → → →
Study registered/first received [registry]	August 24,2005	✘
Study start [registry]	October 2005	✘
Recruitment started [publication]	NR	
Recruitment completed [publication]	NR	
Primary outcome completed [registry]	July 2008	✘
Study completed [registry]	July 2008	
Results first received [registry]	July 21, 2009	✘
Publication submitted [publication]	August 14,2009	✘
Publication accepted [publication]	January 20, 2010	✘
Article first published online [publication]	January 27,2010	✘
Publication printed [publication]	May 2010	✘
Registry last updated [registry]	July 21, 2009	✘
Primary outcome changed [registry]	July 21, 2009	✘

Garber, 2009	**Date**	**Time** → → → →
Study registered/first received [registry]	February 20,2006	✘
Study start [registry]	February 2006	✘
Recruitment started [publication]	NR	
Recruitment completed [publication]	NR	
Primary outcome completed [registry]	November 2008	✘
Study completed [registry]	March 2010	✘
Results first received [registry]	February 23,2010	✘
Publication submitted [publication]	NR	
Publication accepted [publication]	NR	
Article first published online [publication]	September 25,2008	✘
Publication printed [publication]	February 2009	✘
Registry last updated [registry]	March 24,2011	✘
Primary outcome changed [registry]	April 16. 2010	✘

Goldberg, 2005	**Date**	**Time** → → →
Study registered/first received [registry]	May 30,2006	✘
Study start [registry]	September 2000	✘
Recruitment started [publication]	NR	
Recruitment completed [publication]	NR	
Primary outcome completed [registry]	March 2004	✘
Study completed [registry]	March 2004	✘
Results first received [registry]	none posted	
Publication submitted [publication]	February 10,2005	✘
Publication accepted [publication]	March 31,2005	✘
Article first published online [publication]	NA	
Publication printed [publication]	July 2005	✘
Registry last updated [registry]	July 1,2010	✘
Primary outcome changed [registry]	NA	

Goldstein, 2007	**Date**	**Time**	→	→	→	→	→
Study registered/first received [registry]	February 15, 2005	✘					
Study start [registry]	March 2005		✘				
Recruitment started [publication]	NR						
Recruitment completed [publication]	NR						
Primary outcome completed [registry]	July 2006			✘			
Study completed [registry]	February 2008					✘	
Results first received [registry]	February 19, 2009						✘
Publication submitted [publication]	March 30, 2007			✘			
Publication accepted [publication]	May 2, 2007				✘		
Article first published online [publication]	May 7, 2007				✘		
Publication printed [publication]	August 2007					✘	
Registry last updated [registry]	April 7, 2010						✘
Primary outcome changed [registry]	April 7, 2010						✘

Note- trials that reported more dates have wider displays.

Gupta, 2009	**Date**	**Time** → → → →
Study registered/first received [registry]	September 14,2005	✘
Study start [registry]	February 2003	✘
Recruitment started [publication]	NR	
Recruitment completed [publication]	NR	
Primary outcome completed [registry]	December 2006	✘
Study completed [registry]	December 2006	✘
Results first received [registry]	none posted	
Publication submitted [publication]	February 12, 2008	✘
Publication accepted [publication]	July 24, 2008	✘
Article first published online [publication]	October 13,2008	✘
Publication printed [publication]	April 2009	✘
Registry last updated [registry]	February 2,2010	✘
Primary outcome changed [registry]	March 4, 2008	✘

Hamann, 2008	**Date**	**Time** → → →
Study registered/first received [registry]	July 28,2006	✘
Study start [registry]	February 2004	✘
Recruitment started [publication]	NR	
Recruitment completed [publication]	NR	
Primary outcome completed [registry]	NR	
Study completed [registry]	NR	
Results first received [registry]	NR	
Publication submitted [publication]	October 13,2006	✘
Publication accepted [publication]	June 6,2007	✘
Article first published online [publication]	December 20,2007	✘
Publication printed [publication]	January 2008	✘
Registry last updated [registry]	May 15,2009	✘
Primary outcome changed [registry]	NA	

Jadzinsky, 2009	**Date**	**Time** → → → → →
Study registered/first received [registry]	May 15, 2006	✘
Study start [registry]	May 2006	✘
Recruitment started [publication]	May 30, 2006	✘
Recruitment completed [publication]	June 1,2007	✘
Primary outcome completed [registry]	November 2007	✘
Study completed [registry]	December 2008	✘
Results first received [registry]	August 17, 2009	✘
Publication submitted [publication]	January 12, 2009	✘
Publication accepted [publication]	March 2, 2009	✘
Article first published online [publication]	May 6,2009	✘
Publication printed [publication]	June 2009	✘
Registry last updated [registry]	August 4, 2010	✘
Primary outcome changed [registry]	June 30,2010	✘

Kaku, 2009	**Date**	**Time** →	→	→
Study registered/first received [registry]	April 4,2008		✘	
Study start [registry]	April 2005	✘		
Recruitment started [publication]	NR			
Recruitment completed [publication]	NR			
Primary outcome completed [registry]	NR			
Study completed [registry]	October 2006	✘		
Results first received [registry]	August 27,2010			✘
Publication submitted [publication]	NR			
Publication accepted [publication]	February 13, 2009		✘	
Article first published online [publication]	March 23,2009		✘	
Publication printed [publication]	May 2009		✘	
Registry last updated [registry]	August 27,2010			✘
Primary outcome changed [registry]	NA			

Note- trials that reported more dates have wider displays.

Nauck, 2007	**Date**	**Time** → → → → →
Study registered/first received [registry]	October 22,2004	✘
Study start [registry]	September 2004	✘
Recruitment started [publication]	NR	
Recruitment completed [publication]	NR	
Primary outcome completed [registry]	May 2006	✘
Study completed [registry]	May 2007	
Results first received [registry]	September 24,2009	✘
Publication submitted [publication]	October 24,2006	✘
Publication accepted [publication]	December 18, 2006	✘
Article first published online [publication]	Jan 26,2007	✘
Publication printed [publication]	March 2007	✘
Registry last updated [registry]	April 7,2010	✘
Primary outcome changed [registry]	NA	

Nauck, 2009	**Date**	**Time** → → → → →
Study registered/first received [registry]	April 25,2006	✘
Study start [registry]	May 2006	✘
Recruitment started [publication]	NR	
Recruitment completed [publication]	NR	
Primary outcome completed [registry]	May 2007	✘
Study completed [registry]	November 2008	✘
Results first received [registry]	February 23,2010	✘
Publication submitted [publication]	July 22, 2008	✘
Publication accepted [publication]	September 28, 2008	✘
Article first published online [publication]	October 17,2008	✘
Publication printed [publication]	January 2009	✘
Registry last updated [registry]	April 16,2010	✘
Primary outcome changed [registry]	April 16,2010	✘

Perez, 2009	**Date**	**Time** → → → →
Study registered/first received [registry]	July 30,2008	✘
Study start [registry]	June 2007	✘
Recruitment started [publication]	NR	
Recruitment completed [publication]	NR	
Primary outcome completed [registry]	August 2008	✘
Study completed [registry]	August 2008	
Results first received [registry]	August 28,2009	✘
Publication submitted [publication]	NR	
Publication accepted [publication]	September 21, 2009	✘
Article first published online [publication]	October 14,2009	✘
Publication printed [publication]	December 2009	✘
Registry last updated [registry]	July 27,2011	✘
Primary outcome changed [registry]	July 1, 2010	✘

Pratley, 2010	Date	Time → → →
Study registered/first received [registry]	June 18,2008	✘
Study start [registry]	June 2008	✘
Recruitment started [publication]	June 16, 2008	✘
Recruitment completed [publication]	June 11,2009	✘
Primary outcome completed [registry]	June 2009	✘
Study completed [registry]	June 2010	✘
Results first received [registry]	June 11,2010	✘
Publication submitted [publication]	NR	
Publication accepted [publication]	NR	
Article first published online [publication]	April 22,2010	✘
Publication printed [publication]	April 2010	✘
Registry last updated [registry]	September 22, 2011	✘
Primary outcome changed [registry]	August 3,2010	✘

Note- trials that reported more dates have wider displays.

Raskin, 2009	**Date**	**Time** → → → →
Study registered/first received [registry]	November 14,2006	✘
Study start [registry]	November 2006	✘
Recruitment started [publication]	NR	
Recruitment completed [publication]	NR	
Primary outcome completed [registry]	November 2007	✘
Study completed [registry]	November 2007	✘
Results first received [registry]	none posted	
Publication submitted [publication]	December 22, 2008	✘
Publication accepted [publication]	March 10,2009	✘
Article first published online [publication]	May 19,2009	✘
Publication printed [publication]	September 2009	✘
Registry last updated [registry]	September 22, 2011	✘
Primary outcome changed [registry]	June 11, 2009	✘

Raz, 2008	**Date**	**Time** → → → →
Study registered/first received [registry]	June 14,2006	✘
Study start [registry]	June 2006	✘
Recruitment started [publication]	NR	
Recruitment completed [publication]	NR	
Primary outcome completed [registry]	May 2007	✘
Study completed [registry]	Aug 2007	✘
Results first received [registry]	September 24,2009	✘
Publication submitted [publication]	NR	
Publication accepted [publication]	December 14,2007	✘
Article first published online [publication]	January 11,2008	✘
Publication printed [publication]	February 2008	✘
Registry last updated [registry]	May 27,2010	✘
Primary outcome changed [registry]	NA	

Rigby, 2009	**Date**	**Time** → → →
Study registered/first received[registry]	June 7,2007	✘
Study start [registry]	May 2007	✘
Recruitment started [publication]	NR	
Recruitment completed [publication]	NR	
Primary outcome completed [registry]	April 2008	✘
Study completed [registry]	April 2008	✘
Results first received [registry]	April 29,2009	✘
Publication submitted [publication]	NR	
Publication accepted [publication]	NR	
Article first published online [publication]	September 28,2009	✘
Publication printed [publication]	January-February 2010	✘
Registry last updated [registry]	June 17,2009	✘
Primary outcome changed [registry]	NA	

Robbins, 2007	**Date**	**Time**	→	→	→
Study registered/first received [registry]	September 12,2005		✘		
Study start [registry]	December 2003	✘			
Recruitment started [publication]	NR				
Recruitment completed [publication]	NR				
Primary outcome completed [registry]	NR				
Study completed [registry]	September 2005		✘		
Results first received [registry]	NA				
Publication submitted [publication]	NR				
Publication accepted [publication]	August 24, 2007			✘	
Article first published online [publication]	December 23,2007				✘
Publication printed [publication]	November 2007				✘
Registry last updated [registry]	October 12,2010				✘
Primary outcome changed [registry]	NA				

Note- trials that reported more dates have wider displays.

Rosenstock, 2006	**Date**	**Time** → → →
Study registered/first received [registry]	July 9, 2007	✘
Study start [registry]	October 2003	✘
Recruitment started [publication]	NR	
Recruitment completed [publication]	NR	
Primary outcome completed [registry]	NR	
Study completed [registry]	NR	
Results first received [registry]	None posted	
Publication submitted [publication]	June 2,2006	✘
Publication accepted [publication]	August 11,2006	✘
Article first published online [publication]	October 4,2006	✘
Publication printed [publication]	November 2006	✘
Registry last updated [registry]	March 17,2011	✘
Primary outcome changed [registry]	NA	

Scott, 2008	**Date**	**Time** → → → →
Study registered/first received [registry]	October 5,2007	✘
Study start [registry]	June 2006	✘
Recruitment started [publication]	NR	
Recruitment completed [publication]	NR	
Primary outcome completed [registry]	March 2007	✘
Study completed [registry]	March 2007	✘
Results first received [registry]	May 17,2010	✘
Publication submitted [publication]	September 26, 2007	✘
Publication accepted [publication]	November 14, 2007	✘
Article first published online [publication]	January 14,2008	✘
Publication printed [publication]	October 2008	✘
Registry last updated [registry]	December 17,2010	✘
Primary outcome changed [registry]	NA	

Seino, 2010	**Date**	**Time** → → → →
Study registered/first received [registry]	October 27,2006	✘
Study start [registry]	November 2006	✘
Recruitment started [publication]	NR	
Recruitment completed [publication]	NR	
Primary outcome completed [registry]	November 2007	✘
Study completed [registry]	May 2008	✘
Results first received [registry]	February 23, 2010	✘
Publication submitted [publication]	NR	
Publication accepted [publication]	February 3, 2010	✘
Article first published online [publication]	March 3,2010	✘
Publication printed [publication]	May 2010	✘
Registry last updated [registry]	March 29,2010	✘
Primary outcome changed [registry]	NA	

van der Meer, 2009	**Date**	**Time**	→	→	→	
Study registered/first received [registry]	December 20,2005		✘			
Study start [registry]	September 1,2004	✘				
Recruitment started [publication]	NR					
Recruitment completed [publication]	NR					
Primary outcome completed [registry]	NR					
Study completed [registry]	September 1,2006			✘		
Results first received [registry]	NR					
Publication submitted [publication]	July 1, 2008				✘	
Publication accepted [publication]	January 27, 2009					✘
Article first published online [publication]	April 6,2009					✘
Publication printed [publication]	April 2009					✘
Registry last updated [registry]	May 11,2010					✘
Primary outcome changed [registry]	NA					

Note- trials that reported more dates have wider displays.