Cochrane Library

Cochrane Database of Systematic Reviews

Audit and feedback: effects on professional practice and healthcare outcomes (Review)

Ivers N, Jamtvedt G, Flottorp S, Young JM, Odgaard-Jensen J, French SD, O'Brien MA, Johansen M, Grimshaw J, Oxman AD

Ivers N, Jamtvedt G, Flottorp S, Young JM, Odgaard-Jensen J, French SD, O'Brien MA, Johansen M, Grimshaw J, Oxman AD.
Audit and feedback: effects on professional practice and healthcare outcomes.
Cochrane Database of Systematic Reviews 2012, Issue 6. Art. No.: CD000259.
DOI: 10.1002/14651858.CD000259.pub3.

TABLE OF CONTENTS

ABSTRACT 1
PLAIN LANGUAGE SUMMARY 2
SUMMARY OF FINDINGS 4
BACKGROUND 6
OBJECTIVES 6
METHODS 7
RESULTS 10
Figure 1. 11
Figure 2. 13
Figure 3. 15
Figure 4. 16
Figure 5. 17
Figure 6. 18
Figure 7. 19
Figure 8. 20
Figure 9. 21
Figure 10. 22
DISCUSSION 28
AUTHORS' CONCLUSIONS 31
ACKNOWLEDGEMENTS 31
REFERENCES 32
CHARACTERISTICS OF STUDIES 51
ADDITIONAL TABLES 203
APPENDICES 208
WHAT'S NEW 218
HISTORY 219
CONTRIBUTIONS OF AUTHORS 219
DECLARATIONS OF INTEREST 219
SOURCES OF SUPPORT 219
INDEX TERMS 220

[Intervention Review]

Audit and feedback: effects on professional practice and healthcare outcomes

Noah Ivers[1], Gro Jamtvedt[2], Signe Flottorp[2], Jane M Young[3], Jan Odgaard-Jensen[2], Simon D French[4], Mary Ann O'Brien[5], Marit Johansen[6], Jeremy Grimshaw[7,8], Andrew D Oxman[6]

[1]Department of Family Medicine, Women's College Hospital, Toronto, Canada. [2]Norwegian Knowledge Centre for the Health Services, Oslo, Norway. [3]Cancer Epidemiology and Services Research, Sydney School of Public Health, University of Sydney, Camperdown, Australia. [4]Primary Care Research Unit, University of Melbourne, Carlton, Australia. [5]Department of Family and Community Medicine, University of Toronto, Toronto, Canada. [6]Global Health Unit, Norwegian Knowledge Centre for the Health Services, Oslo, Norway. [7]Clinical Epidemiology Program, Ottawa Hospital Research Institute, Ottawa, Canada. [8]Department of Medicine, University of Ottawa, Ottawa, Canada

Contact: Noah Ivers, Department of Family Medicine, Women's College Hospital, 76 Grenville Street, Toronto, ON, M5S1B2, Canada. noah.ivers@utoronto.ca.

Editorial group: Cochrane Effective Practice and Organisation of Care Group.
Publication status and date: Edited (no change to conclusions), published in Issue 7, 2012.

Citation: Ivers N, Jamtvedt G, Flottorp S, Young JM, Odgaard-Jensen J, French SD, O'Brien MA, Johansen M, Grimshaw J, Oxman AD. Audit and feedback: effects on professional practice and healthcare outcomes. *Cochrane Database of Systematic Reviews* 2012, Issue 6. Art. No.: CD000259. DOI: 10.1002/14651858.CD000259.pub3.

ABSTRACT

Background

Audit and feedback is widely used as a strategy to improve professional practice either on its own or as a component of multifaceted quality improvement interventions. This is based on the belief that healthcare professionals are prompted to modify their practice when given performance feedback showing that their clinical practice is inconsistent with a desirable target. Despite its prevalence as a quality improvement strategy, there remains uncertainty regarding both the effectiveness of audit and feedback in improving healthcare practice and the characteristics of audit and feedback that lead to greater impact.

Objectives

To assess the effects of audit and feedback on the practice of healthcare professionals and patient outcomes and to examine factors that may explain variation in the effectiveness of audit and feedback.

Search methods

We searched the Cochrane Central Register of Controlled Trials (CENTRAL) 2010, Issue 4, part of *The Cochrane Library.* www.thecochranelibrary.com, including the Cochrane Effective Practice and Organisation of Care (EPOC) Group Specialised Register (searched 10 December 2010); MEDLINE, Ovid (1950 to November Week 3 2010) (searched 09 December 2010); EMBASE, Ovid (1980 to 2010 Week 48) (searched 09 December 2010); CINAHL, Ebsco (1981 to present) (searched 10 December 2010); Science Citation Index and Social Sciences Citation Index, ISI Web of Science (1975 to present) (searched 12-15 September 2011).

Selection criteria

Randomised trials of audit and feedback (defined as a summary of clinical performance over a specified period of time) that reported objectively measured health professional practice or patient outcomes. In the case of multifaceted interventions, only trials in which audit and feedback was considered the core, essential aspect of at least one intervention arm were included.

Data collection and analysis

All data were abstracted by two independent review authors. For the primary outcome(s) in each study, we calculated the median absolute risk difference (RD) (adjusted for baseline performance) of compliance with desired practice compliance for dichotomous outcomes and the median percent change relative to the control group for continuous outcomes. Across studies the median effect size was weighted by number of health professionals involved in each study. We investigated the following factors as possible explanations for the variation in the effectiveness of interventions across comparisons: format of feedback, source of feedback, frequency of feedback, instructions for improvement, direction of change required, baseline performance, profession of recipient, and risk of bias within the trial itself. We also conducted exploratory analyses to assess the role of context and the targeted clinical behaviour. Quantitative (meta-regression), visual, and qualitative analyses were undertaken to examine variation in effect size related to these factors.

Main results

We included and analysed 140 studies for this review. In the main analyses, a total of 108 comparisons from 70 studies compared any intervention in which audit and feedback was a core, essential component to usual care and evaluated effects on professional practice. After excluding studies at high risk of bias, there were 82 comparisons from 49 studies featuring dichotomous outcomes, and the weighted median adjusted RD was a 4.3% (interquartile range (IQR) 0.5% to 16%) absolute increase in healthcare professionals' compliance with desired practice. Across 26 comparisons from 21 studies with continuous outcomes, the weighted median adjusted percent change relative to control was 1.3% (IQR = 1.3% to 28.9%). For patient outcomes, the weighted median RD was -0.4% (IQR -1.3% to 1.6%) for 12 comparisons from six studies reporting dichotomous outcomes and the weighted median percentage change was 17% (IQR 1.5% to 17%) for eight comparisons from five studies reporting continuous outcomes. Multivariable meta-regression indicated that feedback may be more effective when baseline performance is low, the source is a supervisor or colleague, it is provided more than once, it is delivered in both verbal and written formats, and when it includes both explicit targets and an action plan. In addition, the effect size varied based on the clinical behaviour targeted by the intervention.

Authors' conclusions

Audit and feedback generally leads to small but potentially important improvements in professional practice. The effectiveness of audit and feedback seems to depend on baseline performance and how the feedback is provided. Future studies of audit and feedback should directly compare different ways of providing feedback.

PLAIN LANGUAGE SUMMARY

Audit and feedback: effects on professional practice and patient outcomes

Researchers in The Cochrane Collaboration conducted a review to evaluate the effect of audit and feedback on the behaviour of health professionals and the health of their patients. After searching for all relevant studies, they found 140 studies that met their requirements. Their findings are summarised below.

The use of audit and feedback to influence health professional behaviour and patient health

In an audit and feedback process, an individual's professional practice or performance is measured and then compared to professional standards or targets. In other words, their professional performance is "audited". The results of this comparison are then fed back to the individual. The aim of this process is to encourage the individual to follow professional standards.

Audit and feedback is often used in healthcare organisations to improve health professionals' performance. It is often used together with other interventions, such as educational meetings or reminders. Most of the studies in this review measured the effect of audit and feedback on doctors, although some studies measured the effect on nurses or pharmacists. Audit and feedback was used to influence their performance in different areas, including the proper use of treatments or laboratory tests or improving the overall management of patients with chronic disease such as heart disease or diabetes.

After their performance had been measured, the health professionals were given feedback either verbally, in writing, or both. In some studies, this feedback was given to them by the researchers responsible for the study, while in other studies, feedback was given by supervisors or colleagues, by professional organisations or by someone representing their employer. In some studies, health professionals were given feedback only once, while others were given feedback once a week or once a month.

In some studies, health professionals were simply given information about their performance and how this compared to professional standards or targets. In other studies, health professionals were also given a specific target that they personally were expected to reach, or were given an action plan with suggestions or advice about how to improve their performance.

What happens when health professionals are given audit and feedback?

The effect of using audit and feedback varied widely across the included studies. Overall, the review shows that:

The effect of audit and feedback on professional behaviour and on patient outcomes ranges from little or no effect to a substantial effect. The quality of the evidence is moderate.

Audit and feedback may be most effective when:

1. the health professionals are not performing well to start out with;

2. the person responsible for the audit and feedback is a supervisor or colleague;

3. it is provided more than once;

4. it is given both verbally and in writing;

5. it includes clear targets and an action plan.

In addition, the effect of audit and feedback may be influenced by the type of behaviour it is targeting. It is uncertain whether audit and feedback is more effective when combined with other interventions.

SUMMARY OF FINDINGS

Summary of findings for the main comparison. Summary of findings: Audit and feedback for health professionals

Patient or population: Healthcare professionals

Settings: Primary and secondary care

Intervention: Audit and feedback with or without other interventions[1]

Comparison: Usual care

Outcomes	Absolute improvement[2]	Number of health professionals (studies)	Quality of the evidence (GRADE)	Comments
Compliance with desired practice (dichotomous outcomes)	Median 4.3% absolute increase in desired practice (IQR 0.5% to 16.0%)	82 comparisons from 49 studies.[3] 2310 clusters/groups of health providers (from 32 cluster trials) and 2053 health professionals (from 17 trials allocating individual providers).	⊕⊕⊕⊖ **moderate** [4]	The effect appears to be larger when baseline performance is low, the source is a supervisor or senior colleague, delivered both verbally and written, provided more than once, aims to decrease current behaviours, targets prescribing, and includes both explicit targets and an action plan.
Compliance with desired practice (continuous outcomes)	Median 1.3% improvement in desired practice (IQR 1.3% to 28.9%)	26 comparisons from 21 studies. 661 clusters/groups of health providers (from 13 cluster trials) and 605 health professionals (from 8 trials allocating individual providers).	⊕⊕⊕⊖ **moderate** [4]	
Patient outcomes (dichotomous)	Median percent change -0.4% (IQR -1.3% to 1.6%)	12 comparisons from 6 studies.	⊕⊕⊖⊖ **low** [5]	
Patient outcomes (continuous)	Median percent change 17% (IQR 1.5 to 17%)	8 comparisons from 5 studies.	⊕⊕⊖⊖ **low** [5]	

GRADE Working Group grades of evidence:
High quality: We are very confident that the true effect lies close to that of the estimate of the effect.
Moderate quality: We are moderately confident in the effect estimate: The true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different.
Low quality: Our confidence in the effect estimate is limited: The true effect may be substantially different from the estimate of the effect.
Very low quality: We have very little confidence in the effect estimate: The true effect is likely to be substantially different from the estimate of effect.

1 - The effect of audit and feedback *alone* on professional practice was similar to audit and feedback as the core, essential feature in *multifaceted* interventions.
2 - The post-intervention risk differences are adjusted for pre-intervention differences between the comparison groups to account for baseline differences. The effect was weighted across studies by the number of health professionals involved in the study to ensure that small trials did not contribute as much to the estimate of effect as large trials.
3 - Many studies had more than two arms and therefore contributed multiple comparisons of audit and feedback versus usual care.
4 - We have downgraded the evidence from high to moderate because of inconsistency in the results that could not be fully explained.
5 - We have downgraded the evidence from moderate to low because of the limited number of trials targeting patient outcomes as a primary outcome.

BACKGROUND

Audit and feedback is widely used as a strategy to improve professional practice. This review updates a previous Cochrane review of the effects of audit and feedback (Jamtvedt 2006), where we defined audit and feedback as a 'summary of the clinical performance of healthcare provider(s) over a specified period of time'.

Earlier versions of this review found that audit and feedback can have an effect on professional practice and patient outcomes, but even when it is effective, these effects are generally small to moderate. Furthermore, the impact of audit and feedback is highly variable (Jamtvedt 2003; Jamtvedt 2006; Thomson OBrien 1997a;Thomson OBrien 1997b). While few studies have directly investigated the relative effectiveness of different characteristics of audit and feedback, it does seem that feedback has the greatest effect when baseline compliance with recommended practice was low (Jamtvedt 2006). Due to both the heterogeneity of studies and the methodology of these reviews, we remained limited in our ability to make recommendations regarding characteristics most likely to lead to successful feedback interventions.

Foy et al (Foy 2005) concisely summarised the problem stating that, “Audit and feedback will continue to be an unreliable approach to quality improvement until we learn how and when it works best.”

How the intervention might work

Many theories exist (with multiple overlapping constructs) to further explain how feedback may lead to quality improvement (for a review of such theories, see Grol 2007). Briefly, individual behaviour change theories suggest that feedback may work in many ways, including (but not limited to) changing recipient awareness and beliefs about current practice and subsequent clinical consequences, changing perceived social norms, affecting self-efficacy, or by directing attention to a specific set of tasks (sub-goals). The observation that the effects of audit and feedback are greatest if baseline compliance is low supports the idea that audit and feedback is felt to be effective as a tool to improve practice because it may overcome healthcare providers’ limited ability to self-assess accurately (Davis 2006). Under this assumption, providers are thought to be inherently motivated to improve care, but lacking intention to change their current practices in large part because they are unaware of their suboptimal performance. In turn, they may be prompted to modify their practice if given feedback that their clinical practice was inconsistent with their peers or with accepted guidelines.

Nevertheless, even if intention to change behaviour is strong, the desired action may depend on multiple factors beyond the control of the healthcare provider. Organisational theories focused on quality improvement offer clues regarding potential important effect modifiers, including organisational culture with respect to quality improvement, and the ‘actionability’ of feedback reports (Hysong 2006). Van der Veer et al. (Van der Veer 2010) conducted a systematic review of the impact on quality of care of using medical registries to produce feedback reports to healthcare professionals. They analysed 53 studies of widely varying quality and considered both quantitative and qualitative data. Most of the studies featured multifaceted interventions. They noted that important effect modifiers seemed to be the quality of the data provided to recipients, the motivation and interest of recipients, and the organisational support for quality improvement.

Some potentially important variables are difficult to operationalise in a trial and others have been tested with uncertain results. For instance, although perceived social and professional norms are considered important predictors of behaviour change, there is conflicting evidence regarding the role of peer-comparison in feedback (Kiefe 2001; Søndergaard 2002; Wones 1987). In an attempt to further delineate how to most effectively design and deliver feedback interventions, Hysong (Hysong 2009) completed a re-analysis of the 2006 Cochrane review based on "Feedback Intervention Theory" (Kluger 1996). The results showed greater effectiveness with increasing frequency of the feedback, with written rather than verbal or graphical delivery and with feedback that included information about the correct solution.

Similarly, Gardner and colleagues (Gardner 2010) conducted a re-analysis of the 2006 Cochrane review that applied the Control Theory of Carver and Scheier (Carver 1982), to test target-setting and action plans as effect-modifiers of feedback. Although the results of that re-analysis were inconclusive because very few studies explicitly described their use of targets or action plans, there is empirical evidence from non-health literature to suggest that goal-setting can increase the effectiveness of feedback (Locke 2002), especially if specific and measurable goals are used. However, the role of participant involvement in either target-setting or in feedback interventions seems promising (BMJ 1992) but remains uncertain (Nasser 2008). Other empirical work from the psychology literature has demonstrated the value of action-plans with respect to improving the effectiveness of feedback (Sniehotta 2009).

Regardless of the feedback design, the nature of the clinical change that the feedback tries to encourage may play a role in the effectiveness of the intervention. Qualitative work indicates that it may be easier to comply with guidelines that aim to increase rather than decrease behaviours (Carlsen 2007).

Why it is important to do this review

The aim of the current update is to investigate the effectiveness of audit and feedback to improve processes and outcomes of care and to examine factors that could influence the effectiveness of this intervention. Given the variability in results of the prior review and the inability to satisfactorily explain this based on intuitive factors, this review will attempt to examine multiple theory-informed feedback design characteristics. In so doing, we hope this review will clarify the effectiveness of audit and feedback in general and inform stakeholders regarding how to best employ feedback to change provider behaviours.

OBJECTIVES

We will address three primary questions in this review:

1. Is audit and feedback effective for improving health provider performance and healthcare outcomes?

2. What are the key factors that explain variation in the effectiveness of audit and feedback?

3. How does the effectiveness of audit and feedback compare to other interventions?

For question 1, we considered the following comparisons.

- ***Comparison A.*** Audit and feedback alone or as the core/essential feature of a multifaceted intervention compared with usual care (includes comparisons B and C).
- ***Comparison B.*** Audit and feedback (alone) compared with usual care.
- ***Comparison C.*** Audit and feedback as the core/essential feature of a multifaceted intervention compared with usual care.

For question 2, we considered the following comparisons.

- ***Comparison D.*** Head-to-head comparisons of different types of audit and feedback interventions (effect of changing the way that audit and feedback is designed or delivered).
- ***Comparison E.*** Audit and feedback as the core/essential feature of a multifaceted intervention compared with audit and feedback alone (effect of adding different co-interventions to audit and feedback).

In addition, ***for question 2 we also conducted a meta-regression on the studies in comparison A*** . In the previous review, we subjectively categorised both the "*intensity*" of the feedback intervention and the "*complexity*" of the targeted behaviour, but this approach did not adequately predict feedback effectiveness in a manner that would clearly inform future intervention design. Therefore, to investigate the effectiveness of different ways of providing audit and feedback and other factors that might modify the effects of audit and feedback, studies in this review were characterised according to a selection of variables considered to be both important (based on relevant literature reviewed in the background section above and our knowledge of theories of behaviour change) and accessible in published manuscripts (based on the prior experience of our systematic review authors). Specifically, we used meta-regression to examine the effects of four ways of providing audit and feedback that might increase its effectiveness.

- Providing instruction for improvement with the feedback in the form of specific goals and/or action plans
- Providing verbal feedback in addition to written feedback
- Providing feedback from a senior or respected colleague, supervisor, employer, purchaser or professional standards review organisation (compared with feedback provided by researchers)
- Providing more frequent feedback

We also examined additional factors not related to the intervention itself that might increase the effects of audit and feedback or its apparent effects.

- Lower baseline compliance
- Feedback requiring increasing current behaviours (compared to decreasing behaviours or changing the approach to a clinical problem)
- Audit and feedback targeting health professionals other than physicians
- Higher risk of bias in the primary study

There are many important factors that may predict effectiveness of audit and feedback; the basis for selecting the above factors to examine in a meta-regression and not including other potential effect modifiers is summarised in Appendix 1. (This appendix is not a comprehensive listing of all possible audit-and-feedback questions, but includes the key factors that we considered for inclusion in this update.)

We recognize the importance of context with respect to the effectiveness of an intervention. In particular, the relative complexity of the targeted behaviour likely plays a role in the ability of feedback to increase guideline adherence. To investigate this issue, we conducted a limited number of exploratory subgroup analyses based on the target of the intervention.

For question 3, we considered the following comparison.

- ***Comparison F.*** Audit and feedback alone or as the core/essential feature of a multifaceted intervention compared with other interventions

METHODS

Criteria for considering studies for this review

Types of studies

Randomised controlled trials (RCTs).

Types of participants

Healthcare professionals responsible for patient care. Healthcare professionals in postgraduate training were included, but studies involving only undergraduate students were not.

Types of interventions

Audit and feedback, defined as 'any summary of clinical performance of health care over a specified period of time'. One may alternatively describe an audit and feedback intervention as 'clinical performance feedback'. The feedback may include recommendations for clinical action and may be delivered in a written, electronic or verbal format.

Studies that focused on real-time feedback for procedural skills were excluded as were studies in which the feedback focused on performance on tests or simulated patient interactions. Studies that featured facilitated relay of communication regarding patient status or symptoms but that did not provide a summary of physician performance were also excluded. In general, even if the term 'feedback' was used in the manuscript, the study was excluded if the intervention would be best classified as 'facilitated relay' of patient-specific clinical information or a 'reminder' (especially when the intervention was at the point of care), or any other unique category in the Cochrane Effective Practice and Organisation of Care (EPOC) (EPOC 2002) classification of quality improvement interventions other than 'audit and feedback' (see also: Shojania 2006).

For this update, we only included interventions where we assessed audit and feedback to be a core or essential element. To this end, we categorised studies by the extent to which audit and feedback was the core component of the intervention into three groups: (i) audit and feedback alone; (ii) audit and feedback as a core, essential component of a multifaceted intervention; or (iii) audit and feedback as a component of a multifaceted intervention but not considered 'core and essential'. In multifaceted interventions (which we defined as studies that utilised two or more interventions aiming to change the behaviour of health

professionals), we made the distinction between 'core' and 'not core' by considering whether the other components were likely to be used in the absence of audit and feedback, or whether the audit and feedback seemed to provide the foundation for the rest of the intervention. In cases where the audit and feedback was merely added to a multifaceted intervention that could easily be offered in its absence, the study would be classified as 'not core'.

For comparisons C, D, E, and F, we used the EPOC classification (EPOC 2002) scheme to identify the components of the multifaceted interventions. We used this classification to differentiate between RCTs that tested different ways of designing or delivering an audit and feedback intervention (comparison D) and RCTs that tested whether additional intervention(s) along with audit and feedback were more effective than audit and feedback alone (comparison E). To illustrate, when a suggestion for improvement accompanies the feedback report, it may alternatively be viewed as a co-intervention (clinician education) or as an intrinsic feature of the feedback design (action plan). As with all other abstracted descriptive variables, this process was completed independently by two abstractors and discrepancies resolved through discussion, including other authors as needed.

Types of outcome measures

We focused on objectively measured provider performance in a healthcare setting or patient health outcomes. We abstracted outcomes from the longest available follow-up interval in the original publication, but we did not abstract data from separate articles or companion reports wherein longer term follow-up was reassessed. Studies that provided data only on cost were excluded as were studies that measured knowledge or performance in a test situation only.

Search methods for identification of studies

The current search strategies differ from the strategies used in previous versions of this review. For this version we developed the MEDLINE search strategy based on all MEDLINE indexed and included studies from the previous review versions, in addition to studies known to be eligible for inclusion, but not yet included, a total of 144 records. One hundred and twenty-eight of the 144 records (89%) were identified by the current MEDLINE strategy. We then translated this strategy into the other databases using the appropriate controlled vocabulary as applicable. CENTRAL and CINAHL were searched without time limits. As we searched for RCTs only, MEDLINE was searched from 2005 onwards and EMBASE from 2010 onwards. We expected that MEDLINE records prior to 2005 and EMBASE records prior to 2010 would have been found in CENTRAL. Full search strategies for all databases - for the current update and for the previous review - are available in Appendix 2.

Electronic searches

We searched the following databases:

- Cochrane Central Register of Controlled Trials (CENTRAL) 2010, Issue 4, part of *The Cochrane Library.*www.thecochranelibrary.com, including the Cochrane Effective Practice and Organisation of Care (EPOC) Group Specialised Register (searched 10 December 2010)
- MEDLINE, Ovid (1950 to November Week 3 2010) (searched 09 December 2010)
- EMBASE, Ovid (1980 to 2010 Week 48) (searched 09 December 2010)
- CINAHL, Ebsco (1981 to present) (searched 10 December 2010)
- Science Citation Index and Social Sciences Citation Index, ISI Web of Science (1975 to present) (searched 12-15 September 2011)

Searching other resources

We searched the Science Citation Index and the Social Sciences Citation Index for studies citing all included studies in this review, in addition to selected studies from the review's Additional references list: (Axt-Adam 1993; Balas 1996; Foy 2002; Foy 2005; Gardner 2010; Hysong 2006; Hysong 2009; Van der Veer 2010). Reference lists of all included studies were reviewed and potentially relevant ones are included in the list of Studies awaiting classification, together with potentially relevant studies retrieved from the citation search. These will be included in a future update of this review.

Data collection and analysis

The following methods will be used in updating this review.

Selection of studies

Two review authors (NI, GJ, SFl, or JY) independently screened the titles and abstracts and applied inclusion criteria; complete manuscripts were sought in the case of uncertainty and differences of opinion resolved through consensus. Conference abstracts were included if they provided sufficient data, a full report could be found or missing data could be obtained from the investigators. For this version of the review, we reassessed whether each study from the previous review met the inclusion criteria.

We categorised the extent to which audit and feedback was the 'core' component of the intervention as follows.

- Audit and feedback alone *(included)*
- Audit and feedback as a core, essential component, combined with other interventions categorised according to EPOC classification scheme *(included)*
- Audit and feedback as a component of a multifaceted intervention but not considered ‘core and essential’ *(excluded)*

Multifaceted interventions were defined as including two or more interventions. Where audit and feedback was not considered to be a core, essential component of the intervention, the study was excluded. In other words, this review included multifaceted interventions when the other components were judged to be unlikely to be used in the absence of audit and feedback, or were built around the audit and feedback, which provided the foundation for the rest of the intervention (rather than the audit and feedback being added to a multifaceted intervention that could easily be offered in its absence).

This was assessed independently by two review authors (NI, GJ, SFl, or JY); of all abstracts screened, only eight disagreements regarding inclusion were due to differences in the assessment of whether or not the article was 'core' audit and feedback. All disagreements were resolved by consensus.

Data extraction and management

Data from included studies were abstracted independently by two review authors (NI, GJ, SF, or SFr). A revised version of the EPOC data collection checklist was used to collect information on study design, type of interventions compared, type of targeted behaviour, participants, setting, methods, outcomes, and results. Discrepancies between authors were resolved through discussion. Studies included in the previous review were reassessed due to changes in the data abstraction form and methods for this updated review. For articles included in the previous review, the new variables analysed in this update (instruction for improvement and direction of change required) were abstracted by one author (NI). In all other cases, the variables have been double-abstracted. For numerical results, abstraction was performed by one author (NI) and double-checked by another author (GJ, SF, or SFr).

Assessment of risk of bias in included studies

Two review authors (GJ, NI, SFl, or SFr) independently assessed the risk of bias of each study and extracted data for newly identified studies using a revised data collection form; discrepancies were resolved by consensus with a third author as needed. The risk of bias for each main outcome in all studies included in the review was assessed according to the revised EPOC criteria. The degree of confidence in the estimate of effect across studies was assessed using GRADEpro and the GRADE approach (Guyatt 2008; Schunemann 2008, Schunemann 2009).

An overall assessment of the risk of bias (high, moderate or low risk of bias) was assigned to each of the included studies using the approach suggested in the Chapter 8 of the *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins 2008). Studies with low risk of bias for all key domains or where it seems unlikely for bias to seriously alter the results were considered to have a low risk of bias. Studies where risk of bias in at least one domain was unclear or judged to have some bias that could plausibly raise doubts about the conclusions were considered to have an unclear risk of bias. Studies with a high risk of bias in at least one domain or judged to have serious bias that decreased the certainty of the conclusions were considered to have a high risk of bias. For the studies included in the previous review, one review author (NI) updated the 'Risk of bias' assessment using this approach. Any discrepancies between the conclusions regarding risk of bias using the new and the previous approach were discussed with other review authors and resolved through consensus

Measures of treatment effect

All outcomes were expressed as compliance with desired practice. Professional and patient outcomes were analysed separately. For trials reporting summary and individual measures of performance, the summary measures were used. When several outcomes were reported in a trial we only extracted results for the variable(s) explicitly described as the primary outcome(s). When the primary outcome was not specified we took the variable(s) described in the sample size calculation as the primary outcome. When the primary outcome was still unclear or when the manuscript described several primary outcomes, we calculated the median value across multiple outcomes.

Since important baseline differences between intervention and control groups are frequently found in cluster-randomised trials, our primary analyses was based on estimates of effect that were adjusted for baseline differences. Therefore, only studies providing data on baseline performance were included in the statistical analysis. Baseline compliance, defined as compliance with desired practice (or with the targeted behaviours) prior to the intervention, was treated as a continuous variable ranging from zero to 100%, based on the median value of pre-intervention level of compliance in the audit and feedback group and control group. For dichotomous outcomes, we calculated the adjusted risk difference (RD) as the difference in adherence after the intervention minus the difference before the intervention. A positive RD indicates that performance improved more in the audit and feedback group than in the control group (eg. an adjusted RD of 0.09 indicates an absolute improvement in compliance with targeted behaviours of 9%). For continuous outcomes, we calculated adjusted change relative to the control group as the post-intervention difference in means minus the baseline difference in means divided by the baseline control group mean. As with the adjusted RD, a positive change indicates that performance improved more in the audit and feedback group than in the control group. This is a relative effect rather than an absolute effect; the effect size reflects the baseline performance as well as the change in performance and it is not bound between -100 and +100%.

Unit of analysis issues

Cluster-randomised trials

Due to the nature of the intervention, we expected that most of the trials would be randomised by cluster. Under such circumstances it is necessary to adjust results from primary trials for clustering before they are included in a meta-analyses in order to avoid underestimating the standard error (SE) of the estimate of effect. As in the previous versions of this review, we have not abstracted the observed SEs, P values, or confidence intervals for our statistical analysis, instead performing meta-regression using the number of health professionals as the basis for weighting.

Studies with more than two arms

If more than one comparison from a study with more than two arms were eligible for the same comparison, we adjusted the number of healthcare professionals to avoid double counting. The adjustment was done by dividing the number of healthcare professionals in the shared arm approximately evenly among the comparisons.

Dealing with missing data

Only studies reporting baseline data for primary outcomes were included in the statistical analysis because the previous review identified baseline performance as an important predictor of feedback effectiveness. Missing data regarding the characteristics of the studies or of the audit and feedback intervention were not imputed.

Assessment of heterogeneity

We explored heterogeneity visually by preparing tables, bubble plots and box plots (displaying medians, inter-quartile ranges, and ranges) to explore the size of the observed effects in relationship to each of these variables. The size of the bubble for each comparison corresponds to the number of healthcare professionals who participated. We also plotted the lines from the weighted regression to aid the visual analysis of the bubble plots.

Data synthesis

Across studies, the median effect size was weighted by the number of health professionals involved in the trial reported to ensure that very small trials did not contribute the same to the overall estimate as larger trials. If the number of health professionals was not reported, the number of practices/hospitals/communities was used instead. Thus, the summary statistics in the meta-analyses reported as weighted median adjusted RD or weighted median adjusted change relative to baseline control are weighted by the number of health professionals, while the results reported from individual studies are not. The primary analyses excluded studies at high risk of bias.

Subgroup analysis and investigation of heterogeneity

Visual analyses were supplemented with meta-regression to examine how the size of the effect (adjusted RD) was related to the potential explanatory variables (listed below), weighted according to the number of healthcare professionals. We accounted for baseline differences in compliance by using adjusted estimates of effect to avoid the effect of potentially important baseline differences in compliance between groups. We conducted a multivariable linear regression using main effects only; baseline compliance treated as a continuous explanatory variable and the others as categorical. For this analysis we excluded studies with a high risk of bias. The analyses were conducted using the GLIMMIX procedure in SAS (Version 9.2. SAS Institute Inc., Cary, NC, USA), where we also took the dependency between comparisons from the same trial into account. P values were based on the classical sandwich estimator.

Each comparison was characterised relative to the other variables in the tables, looking at one potential explanatory variable at a time in univariate analyses. If the number of included studies was large enough, we also performed a multivariate analysis including all potential explanatory variables. We assessed the following potential sources of heterogeneity to explain variation in the results of the included studies.

- Format (verbal; written; both; unclear)
- Source (supervisor or senior colleague; professional standards review organisation or representative of employer/purchaser; investigators; unclear)
- Frequency (weekly; monthly; less than monthly; one-time)
- Instruction for improvement (explicit measurable target or specific goal but no action plan; action plan with suggestions or advice given to help participants improve but no goal/target; both; neither)
- Direction of change required (increase current behaviour; decrease current behaviour; mix or unclear)
- Recipient (physician; other health professional)
- Risk of bias (high; unclear; low)
- Baseline compliance (continuous measure of health professionals' compliance with desired practice)

We hypothesised that audit and feedback with the following characteristics would be most effective: provided in both verbal and written format, from a supervisor or senior colleague, delivered more frequently than less, featuring both specific goals and action plans, aiming to increase rather than decrease behaviours, and received by non-physician providers. We also hypothesised that studies with low risk of bias would be associated with smaller effect sizes.

In addition, we conducted two exploratory analyses to examine the importance of context and the relative complexity of the targeted behaviour on the likelihood that feedback would improve professional practice. We compared the effectiveness of feedback in outpatient (primary care or outpatient clinics) and hospital (inpatient, emergency room or hospital) settings. In addition, we considered common targets of feedback interventions, including: appropriate prescribing, test-ordering (laboratory or radiology), and diabetes or cardiovascular disease management (two chronic clinical conditions with similar management and targets). We did not have any a priori hypotheses for these analyses. However, the second analysis reflects two hypotheses that we tested in the previous update of this review: that the effectiveness of feedback would be greater for behaviours that are important but not complex (ie. prescribing) compared to more complex behaviours (ie. disease management) or compared to behaviours that clinicians might perceive as less important (ie. test-ordering). For these analyses, we compared the weighted median effect sizes and conducted a univariate meta-regression for studies reporting dichotomous outcomes. If we found potentially important and statistically significant differences, we included these explanatory factors in the full model for the meta-regression described above to assess the robustness of these exploratory findings.

Sensitivity analysis

We performed sensitivity analyses by including studies with a high risk of bias. We also examined whether differences in the level of the unit of *analysis* (groups of professionals versus individual professionals versus patients) was a source of heterogeneity, since analyses conducted at different levels can result in different effect estimates.

RESULTS

Description of studies

For this update we screened 3623 new studies and reviewed the full text of 282. The total number of studies included is 140. Of note, 53 new studies were added to this review since the previous update and 31 were removed from the previous version of the review as they no longer met our inclusion criteria. See study flow diagram for details (Figure 1).

Figure 1. Study flow diagram.

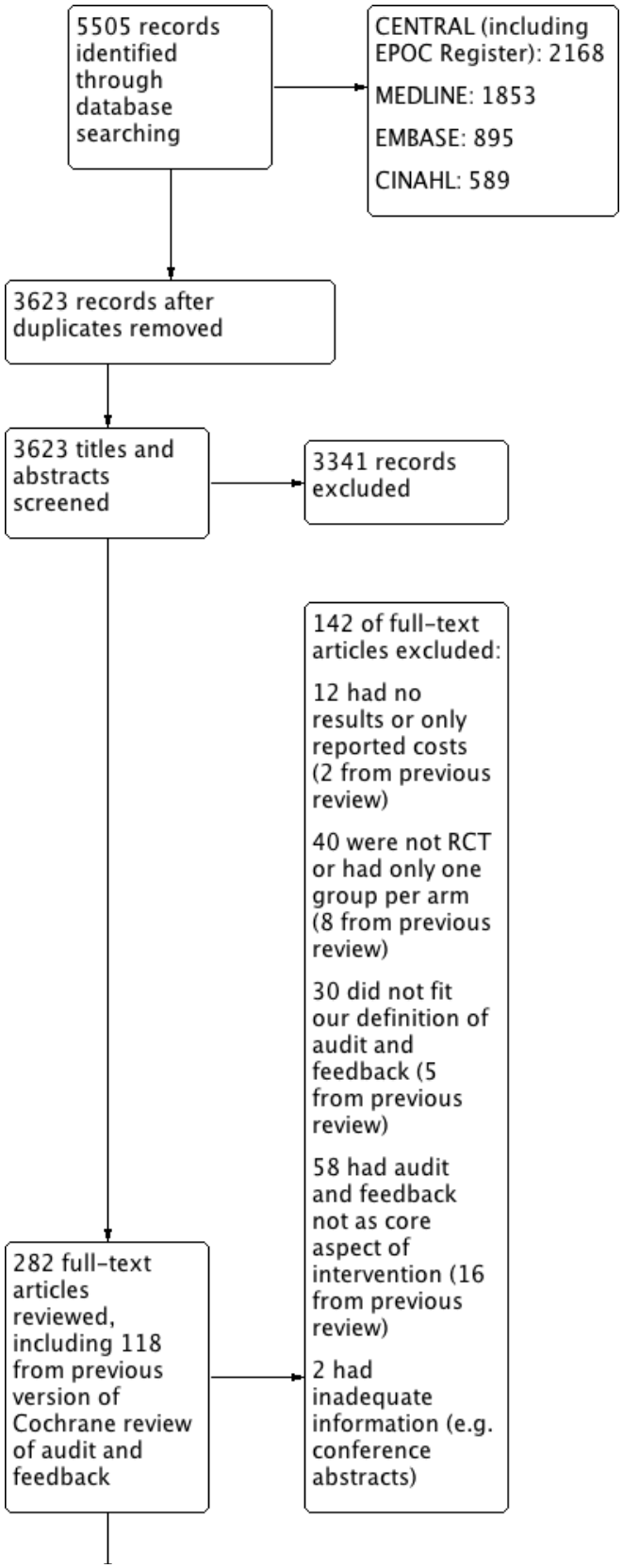
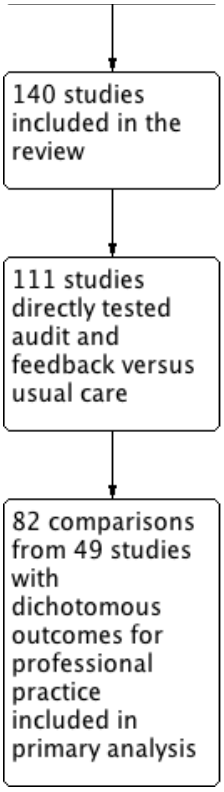


Figure 1. (Continued)



All abstracted information is available upon request; the general characteristics of the included studies are described in Table 1.

The unit of allocation was a single healthcare provider in 51 studies (5056 total providers, median 56), groups of clusters of healthcare professionals (e.g. clinics, wards, hospitals, communities) in 88 studies (5267 total clusters, median 32), and in one study (24 providers, 1140 patients) the unit of allocation was not clear (Everett 1983). Twenty studies had four arms, 22 studies had three and the remaining 98 had two arms.

Characteristics of setting and professionals

Eighty trials were based in North America (69 in USA, 11 in Canada), 21 in the UK or Ireland, 10 in Australia or New Zealand, and 29 elsewhere. Only four studies were from low- and middle-income countries (two in Sudan, one in Thailand, and one in Laos).

In 121 trials the targeted health professionals for the intervention were physicians. Five studies explicitly targeted pharmacists and 16 studies explicitly targeted nurses. The most common clinical specialty area was general or family practice, targeted in 84 trials. Ninety-four trials were in an outpatient setting, 36 were in inpatient settings, and in 10 studies the clinical setting was unclear.

Targeted behaviours

There were 39 trials specifically aiming to improve appropriate prescribing and 31 specifically targeting laboratory or radiology test utilisation. Thirty-four trials focused on management of patients with either cardiovascular disease or diabetes (two exemplar chronic conditions with common management strategies). The remaining trials varied widely across conditions and targeted behaviours.

Characteristics of interventions

There were 49 studies in which audit and feedback was the only intervention, while audit and feedback was considered the core, essential component of a multifaceted intervention in 91 studies.

The format of the feedback was clearly reported in 129 studies: 13 had verbal feedback, 84 had written feedback, and 32 had both. In the majority of studies (112), the source of the feedback was unclear or it was provided by the researchers who had no other relationship to the recipients. In 13 studies feedback was provided from a supervisor or senior colleague, and in 15 from a 'professional standards review organisation' or representative of the employer or purchaser. The frequency of the feedback was weekly in 11 trials, monthly in 19 trials, repeated but less than monthly in 36, and once only in 68 trials.

In 11 studies the feedback provided recipients with explicit, measurable goals and 41 studies included action plans or correct solution information with the feedback. The feedback had both these features in four studies and neither in 84 studies. In 57 studies, the feedback required recipients to increase current behaviours; in 29 they had to decrease current behaviours, and in 55 studies the feedback was judged to require a complex or uncertain change in behaviour.

Outcome measures

There was large variation in outcome measures, and studies often reported multiple primary outcomes related to compliance with

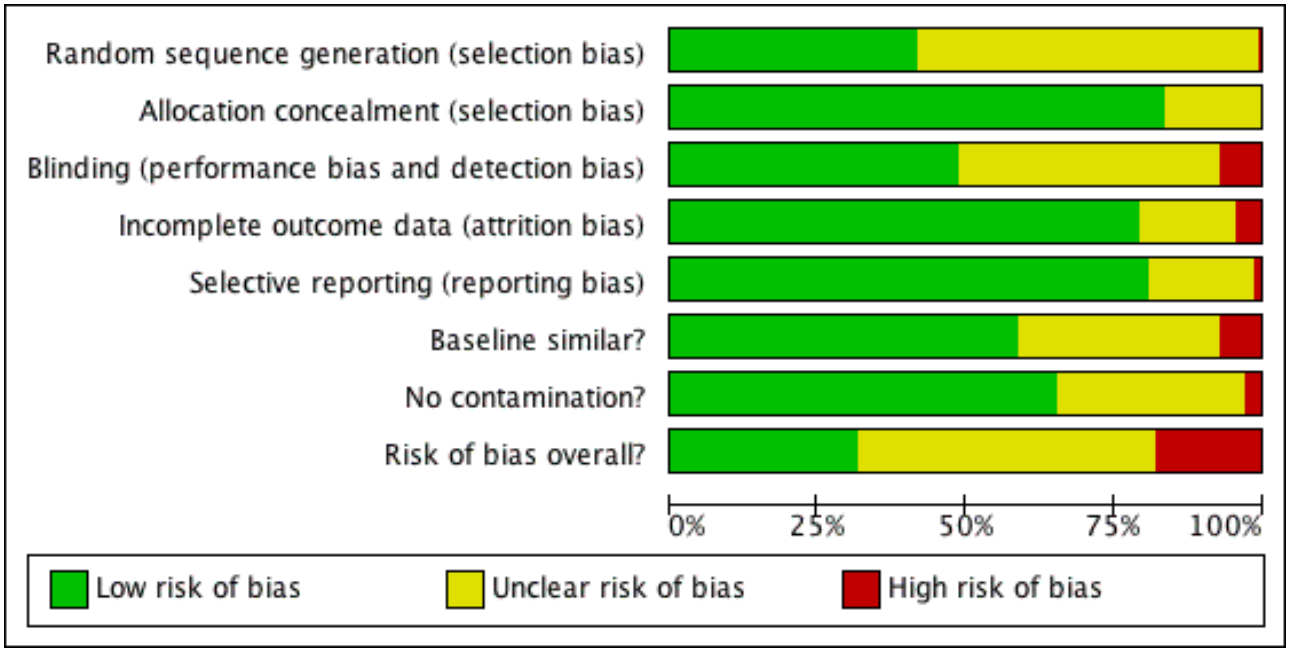
different aspects of a guideline. Most trials measured professional practice, such as prescribing or use of laboratory tests. Some trials reported both practice and patient outcomes such as smoking status or blood pressure. There was a mixture of dichotomous outcomes (for example the proportion compliance with guidelines or the proportion of patients with appropriate management) and continuous outcome measures (for example costs, number of laboratory tests, or number of prescriptions) across and within studies.

Baseline performance was not reported in 10 studies (Balas 1998; Berman 1998; Curtis 2007; Everett 1983; Linn 1980; Lobach 1996; Robling 2002; Sandbaek 1999; Tierney 1986; Wones 1987).

Risk of bias in included studies

See Figure 2. Of the 140 trials, 44 (31%) had a low risk of bias, 71 (51%) had an unclear risk of bias, and 25 (18%) had a high risk of bias (Baker 1997; Batty 2001; Berman 1998; Boekeloo 1990; Brown 1994; Buffington 1991; Canovas 2009; Charrier 2008; Claes 2005; Curran 2008; Everett 1983; Foster 2007; Gama 1992; Gehlbach 1984; Kim 1999; Millard 2008; Robling 2002; Rust 1999; Sandbaek 1999; Schneider 2008; Sommers 1984; Søndergaard 2006; Wadland 2007; Winkens 1995; Zwar 1999). The most common sources of a high risk of bias related to lack of similarity at baseline (ten trials), lack of outcome blinding (e.g. when outcomes were reported by participating healthcare professionals) (ten trials), and due to incomplete follow-up (six trials). Clarity of reporting regarding the risk of bias variables was frequently inadequate. For example, the nature of the randomisation sequence was unclear in 81 trials, outcome blinding was unclear in 61 trials, similarity at baseline was unclear in 48 trials, and risk of contamination was unclear in 45 trials. Randomisation was clearly concealed (or there was cluster randomisation) in 117 trials. There was adequate follow-up in 111 trials.

Figure 2. Risk of bias graph: review authors' judgements about each risk of bias item presented as percentages across all included studies.



Effects of interventions

See: **Summary of findings for the main comparison** Summary of findings: Audit and feedback for health professionals

See Summary of findings for the main comparison.

Comparison A. Any intervention in which audit and feedback is the single intervention or is the core, essential feature of a multifaceted intervention, compared to usual care

A total of 171 comparisons from 109 studies were included in this comparison. Of these, 17 comparisons from 10 studies had no baseline data, and 21 comparisons from 14 studies were at high risk of bias. Twenty-five comparisons from 15 studies included patient outcomes as a primary outcome. Thus, 108 comparisons from 70 studies were included in the primary analyses assessing the effects of audit and feedback on professional practice.

Dichotomous measures of compliance with desired practice

There were 124 total comparisons, of which 11 comparisons were removed due to lack of adequate baseline data. Of the 113 remaining comparisons, 15 had patient-oriented outcomes, leaving 98 comparisons from 62 studies. In the primary meta-analysis, a further 16 comparisons from 12 studies at high risk of bias were excluded, leaving 82 comparisons from 49 studies with dichotomous outcomes. These studies included 2310 clusters/groups of health providers (from 32 cluster trials), and 2053 health professionals (from 17 trials allocating individual providers).

For these studies, the weighted median adjusted RD was a 4.3% increase in compliance with desired practice (interquartile range (IQR) 0.5% to 16%). The weighted median RD when studies with high risk of bias were included in the sensitivity analysis was also 4.3% (IQR 0.6% to 16%).

The range in adjusted RDs for compliance with desired practice was wide: a 9% absolute decrease to a 70% increase in compliance. Of the 98 total comparisons, 27 had an adjusted RD of at least 10% and in 20 comparisons the adjusted RD was between 5% and 10%. For 50 comparisons the adjusted RD was small (ranging from -5% to 5%). Only one study reported a negative effect greater than 5%; an adjusted RD of -9% for appropriate prescribing of benzodiazepines (Batty 2001). This study had a high risk of bias due to imbalance at baseline. Three other studies had unusually large effect sizes. Foster 2007 reported a 45% increase in the utilisation of peak flow in asthma patients. This study had a high risk of bias due to incomplete follow-up. Gehlbach 1984 reported a 45% improvement in the use of generic prescriptions and this study also had a high risk of bias. Finally, Mayer 1998 showed a 70% increase in the provision of skin cancer preventive advice among pharmacists, from a baseline performance of 0%. As in the previous version of this review, this study was excluded from the primary analysis because it differed from the others, as it aimed to initiate an entirely new clinical behaviour in the intervention group, rather than help providers to improve their performance in an area of known professional responsibility.

There were 11 comparisons from seven studies with dichotomous outcomes that did not report baseline data (Balas 1998; Berman 1998; Curtis 2007; Lobach 1996; Robling 2002; Sandbaek 1999; Tierney 1986). The range of (unadjusted) RD seen in these studies was -2.3% to 29.2%. The median unadjusted RD for these studies was 4% (IQR 1% to 7%).

Continuous measures of compliance with desired practice

There were 47 total comparisons, of which six were removed due to lack of adequate baseline data. Of the 41 remaining comparisons with continuous primary outcomes, 10 had patient-oriented outcomes, leaving 31 comparisons from 25 studies. The primary meta-analysis excluded a further five comparisons from four studies at high risk of bias leaving 26 comparisons from 21 studies with continuous outcomes. These studies included 661 groups of healthcare providers (from 13 cluster trials) and 605 healthcare professionals (from eight trials allocating individual providers).

For these studies, the weighted median adjusted change relative to baseline control was a 1.3% increase in compliance with desired practice (IQR 1.3% to 23.2%). When studies at high risk of bias were included, the weighted median adjusted change relative to baseline control was 2.9% (IQR 1.3% to 26.1%).

The adjusted change relative to baseline control varied widely, from a 50% decrease in desired practice to a 139% increase in desired practice. Of the 31 total comparisons with continuous outcomes, 21 had an adjusted change relative to baseline control of at least 10%. For eight comparisons the adjusted change relative to baseline control was relatively small (-5% to 5%). Two comparisons had larger negative effects: one (Holm 1990) showed a 10% relative increase in benzodiazepine/sedative medications; the other comparison (Cohen 1982) showed a 50% relative increase in laboratory test utilisation, but actually reported a positive effect during the intervention period, which reversed after the intervention stopped. The trial (Wadland 2007) that reported a 139% relative increase in smoking cessation referrals had a high risk of bias.

There were six comparisons from three studies with continuous outcomes that did not report baseline data (Everett 1983; Linn 1980; Wones 1987). The median effect seen in these studies was a 54% relative increase in desired practice (IQR 15.1% to 54%)

Patient outcomes

Fifteen studies (Buffington 1991; Curran 2008; Fairbrother 1999; Gullion 1988; Hemminiki 1992; Hendryx 1998; Linn 1980; Lomas 1991; Mitchell 2005; O'Connor 2009; Phillips 2005; Rantz 2001; Rust 1999; Svetkey 2009; Thomas 2007) reported patient-type outcomes as a primary outcome. One study (Linn 1980) did not have any baseline data, and two studies (Buffington 1991; Curran 2008) had a high risk of bias, leaving 12 comparisons with dichotomous outcomes and eight comparisons with continuous outcomes for analysis.

There was minimal discernable effect observed for patient outcomes with dichotomous outcomes, while a positive effect was noted in studies with continuous outcomes. Specifically, for dichotomous outcomes, the weighted median adjusted RD was a 0.4% decrease in desired outcomes (IQR -1.3% to 1.6%) and for continuous outcomes, the weighted median adjusted change relative to baseline control was a 17% improvement (IQR 1.5% to 17%).

Investigation of heterogeneity

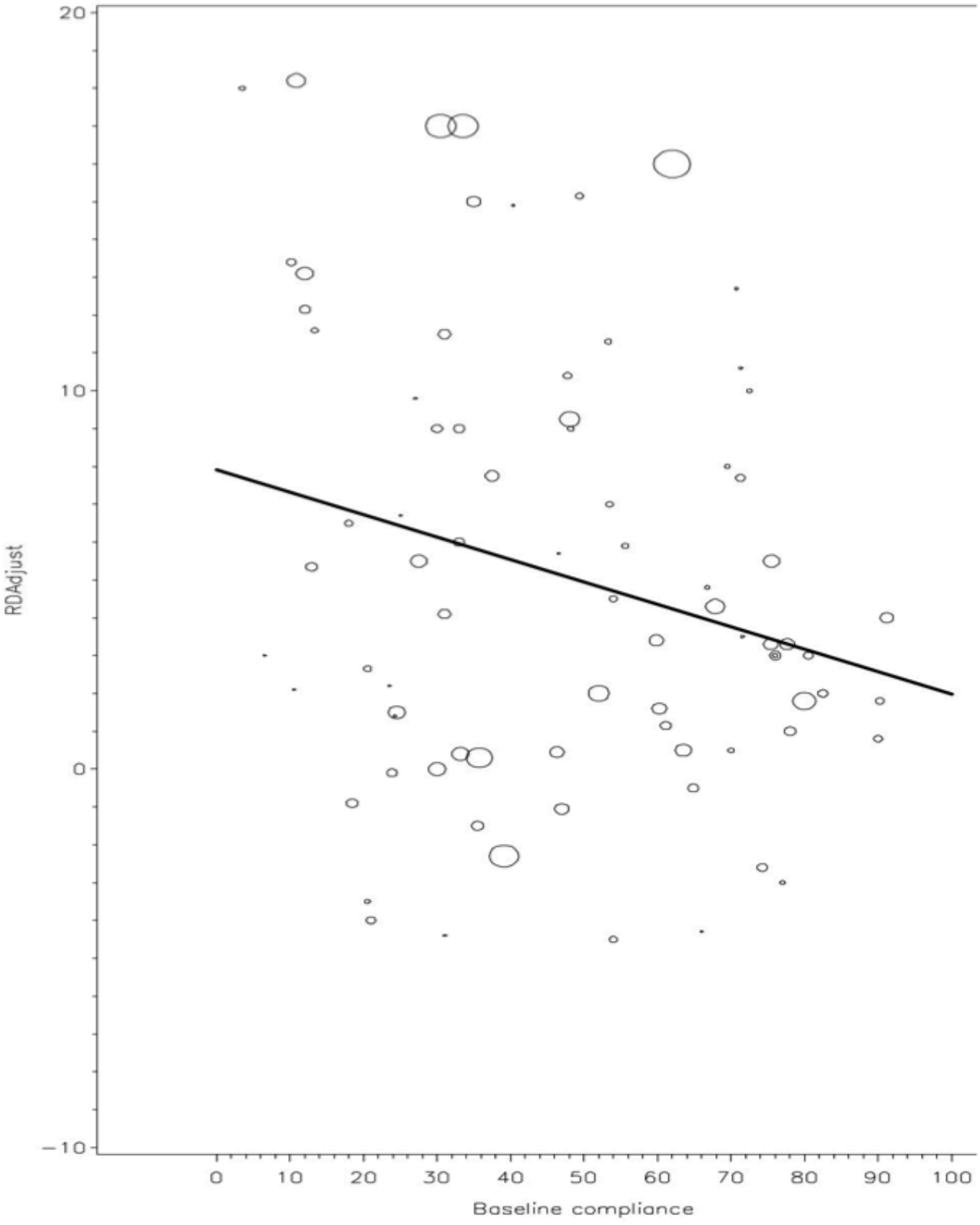
The multivariable meta-regression analysis explored the role of five characteristics of the intervention (format, source, frequency, instructions for improvement, direction of change required), two characteristics of the recipients (baseline performance, profession), and one characteristic of the trial design (risk of bias) on heterogeneity in effect size. This was performed on trials that had dichotomous outcomes and that compared audit and feedback as the only intervention or as the core, essential feature of a multi-faceted intervention versus usual care. Studies at high risk of bias were excluded, leaving 80 comparisons in this analysis with either unclear or low risk of bias.

All five characteristics of the intervention were identified as significant in the model, as described in Table 2, indicating that the format ($P = 0.02$), source ($P < 0.001$), frequency ($P < 0.001$), instructions for improvement ($P < 0.001$), and the direction of change required ($P = 0.007$) each help explain variation in effects. Within these variables, relatively large differences in effect size were seen when comparing certain characteristics: presented in both verbal and written format versus only verbal (expected difference in adjusted RD = 8%); delivered by a supervisor or senior colleague versus the investigators (expected difference in adjusted RD = 11%); frequency of monthly versus once only (expected difference in adjusted RD = 7%); containing both an explicit, measurable target and a specific action plan versus neither (expected difference in adjusted RD = 5%); and requiring a decrease versus an increase of current behaviour to achieve a higher score (expected difference in adjusted RD = 6%).

Risk of bias ($P = 0.679$) and profession (physician versus non-physician) ($P = 0.561$) were not associated with variation in effect size. Lower baseline performance was associated with greater effectiveness for the intervention ($P = 0.007$). To illustrate, the model predicts that recipients who achieved 25% of desired practice at baseline would have an expected adjusted RD of 9%, while those who achieved 75% of desired practice at baseline would

have an expected adjusted RD of only 5%. See Figure 3 for a bubble plot of effect size by baseline performance.

Figure 3. Bubble plot: adjusted risk difference by baseline performance



Examination of box plots for each of the explanatory variables primary analysis supported the statistical conclusions (see Figure 4, Figure 5, Figure 6, Figure 7, Figure 8, Figure 9). For exploratory purposes, we also examined box plots for explanatory variables considering trials with continuous outcomes from Comparison A. This did not result in any qualitative differences in the assessment of heterogeneity. Finally, we examined the box plots for trials with dichotomous and continuous outcomes, respectively, for Comparison B (audit and feedback alone versus usual care) and then for Comparison C (audit and feedback as the core, essential feature of a multifaceted intervention versus usual care), separately. These analyses revealed consistency in the direction of effects for the explanatory variables, supporting the initial conclusions.

Figure 4. Box plot: comparing adjusted risk difference by format of feedback

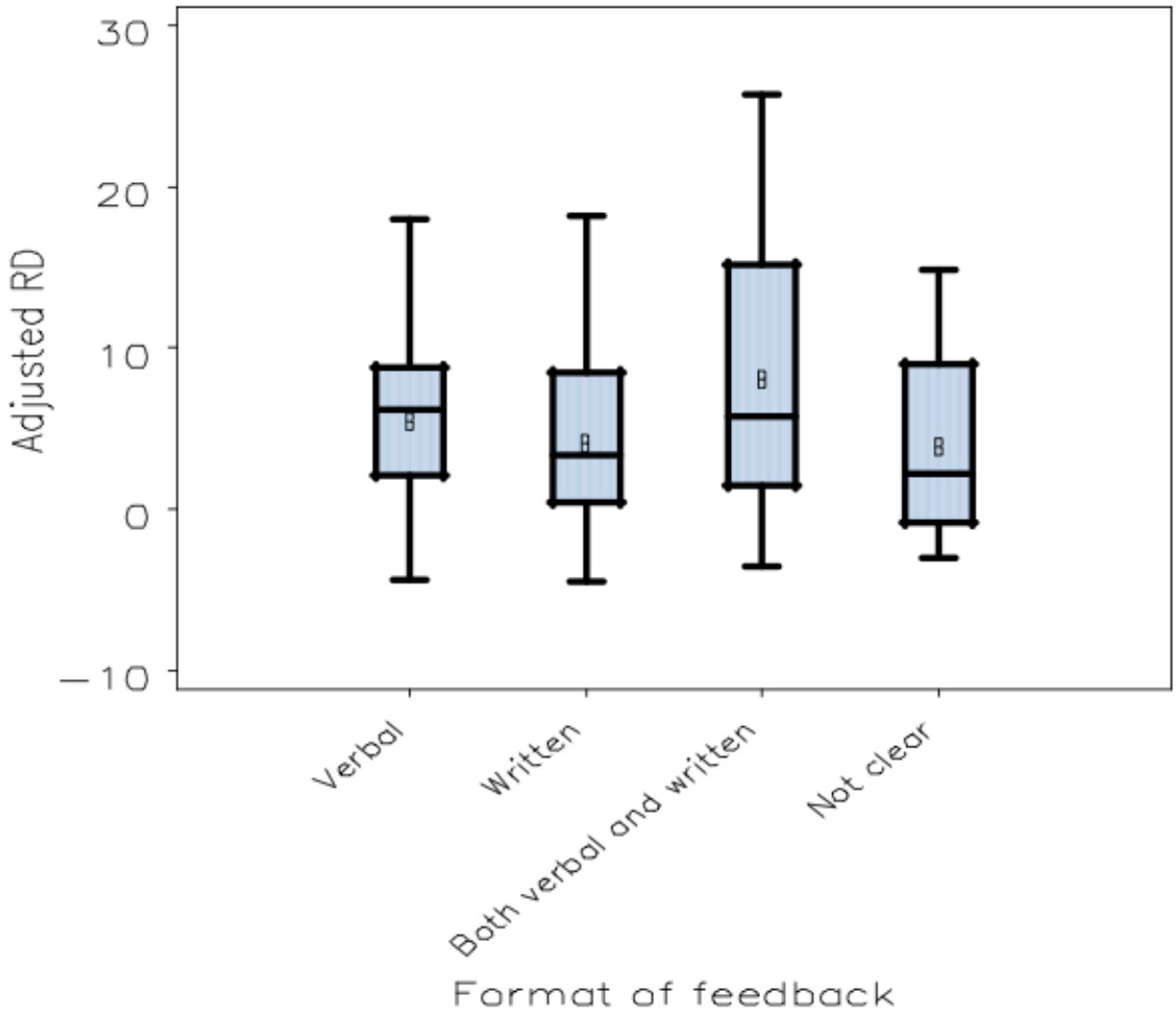


Figure 5. Box plot: comparing adjusted risk difference by format of feedback

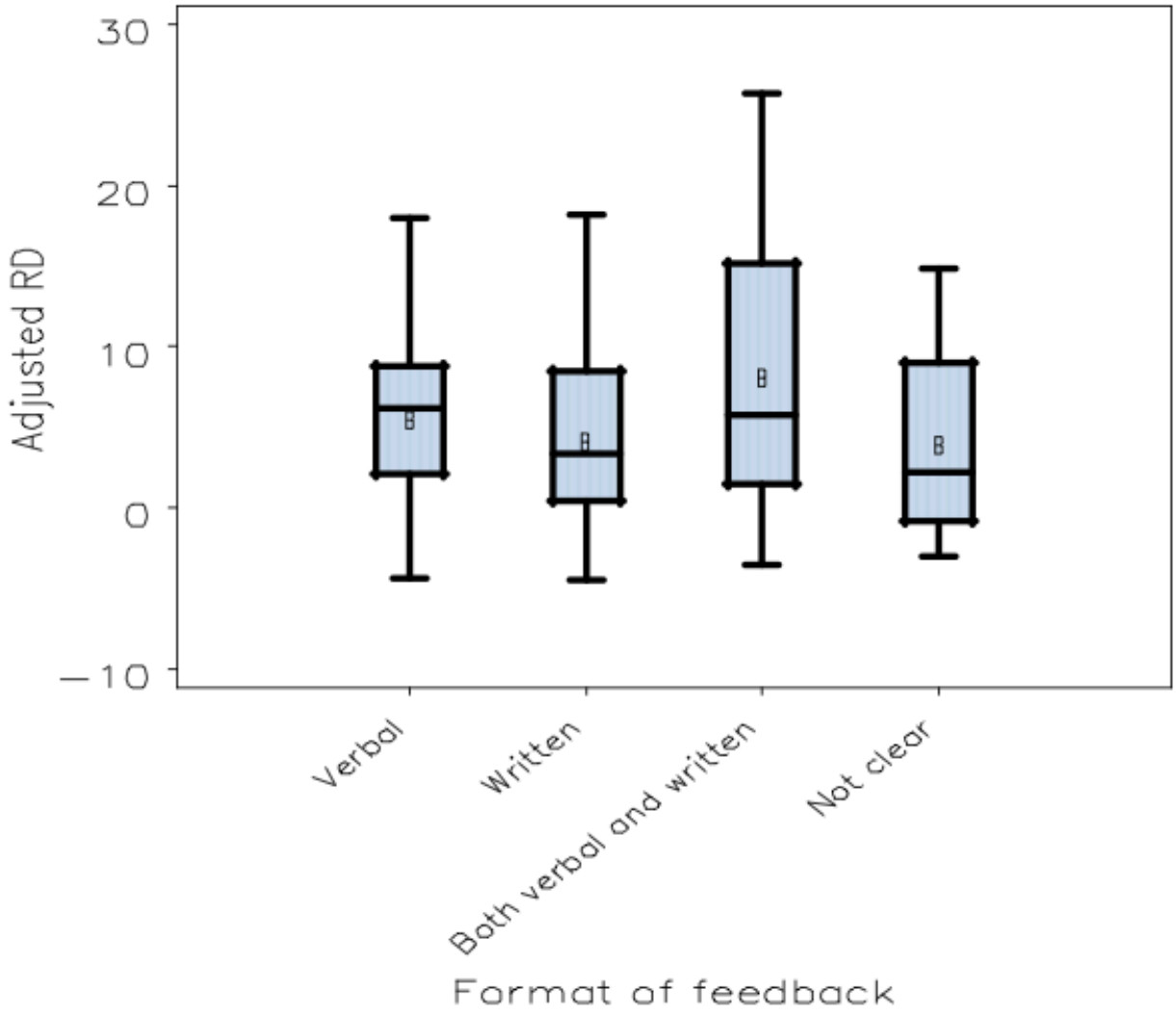


Figure 6. Box plot: comparing adjusted risk difference by source of feedback

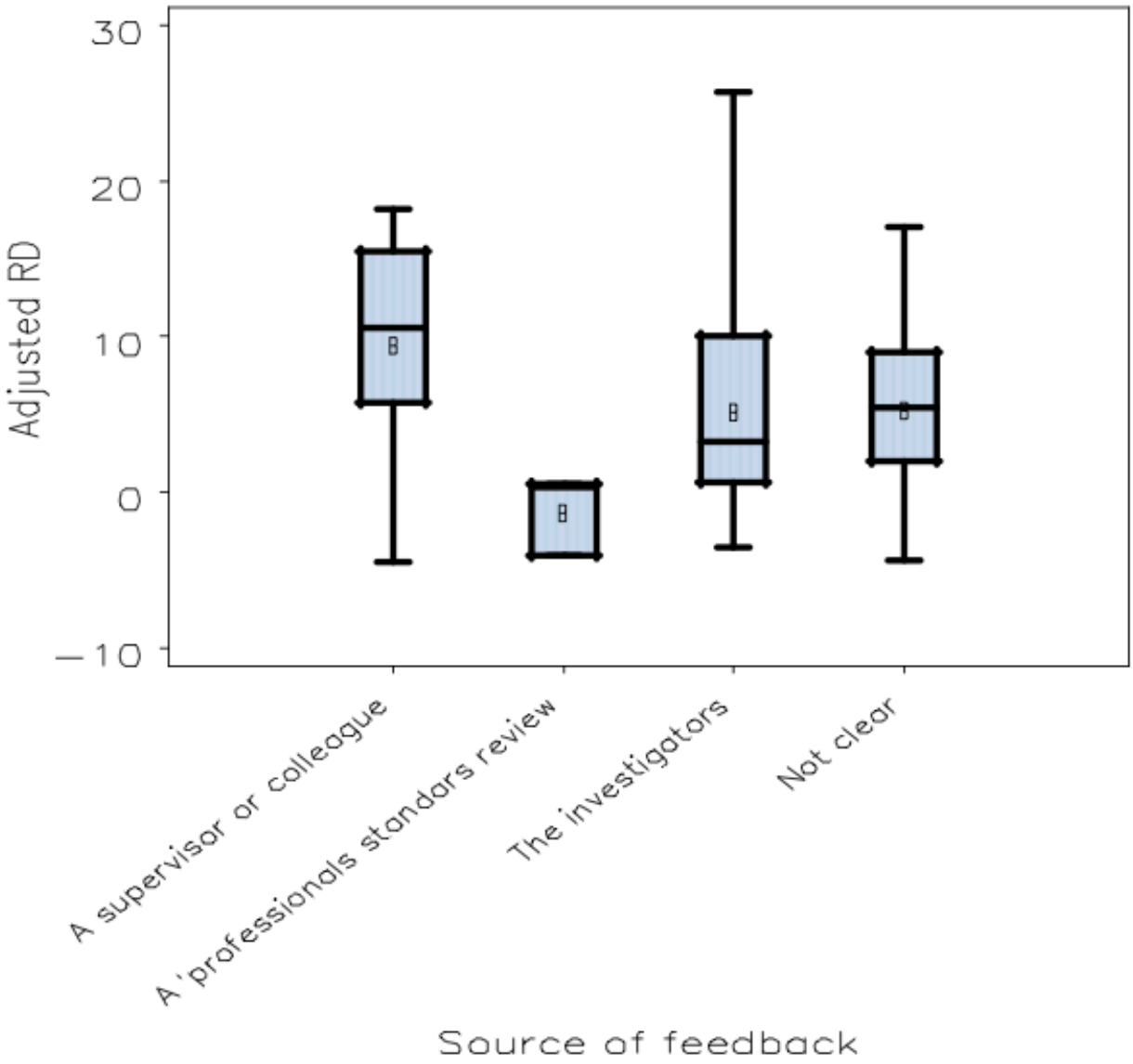


Figure 7. Box plot: comparing adjusted risk difference by frequency of feedback

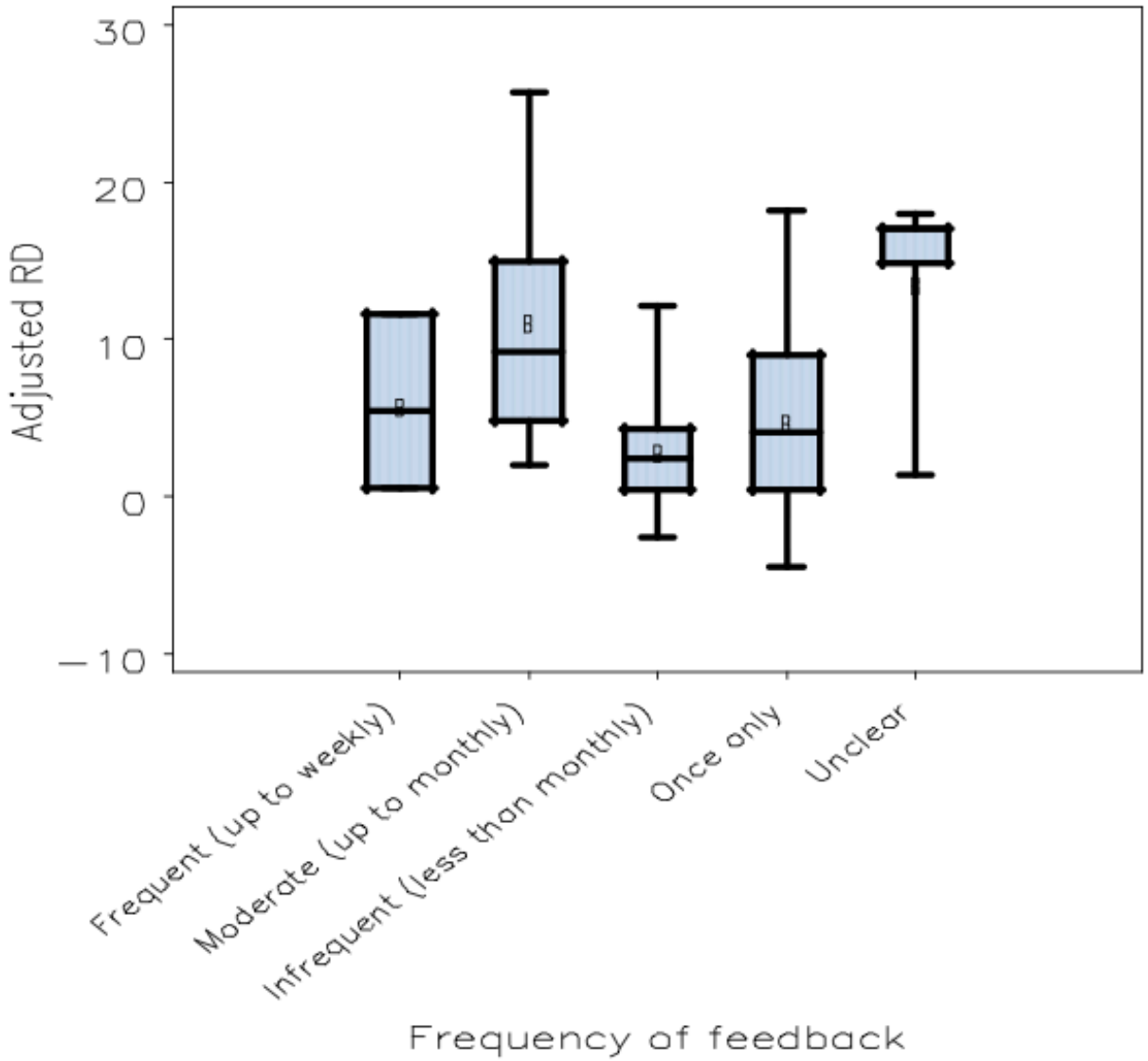


Figure 8. Box plot: comparing adjusted risk difference by presence/extent of instructions for improvement in feedback

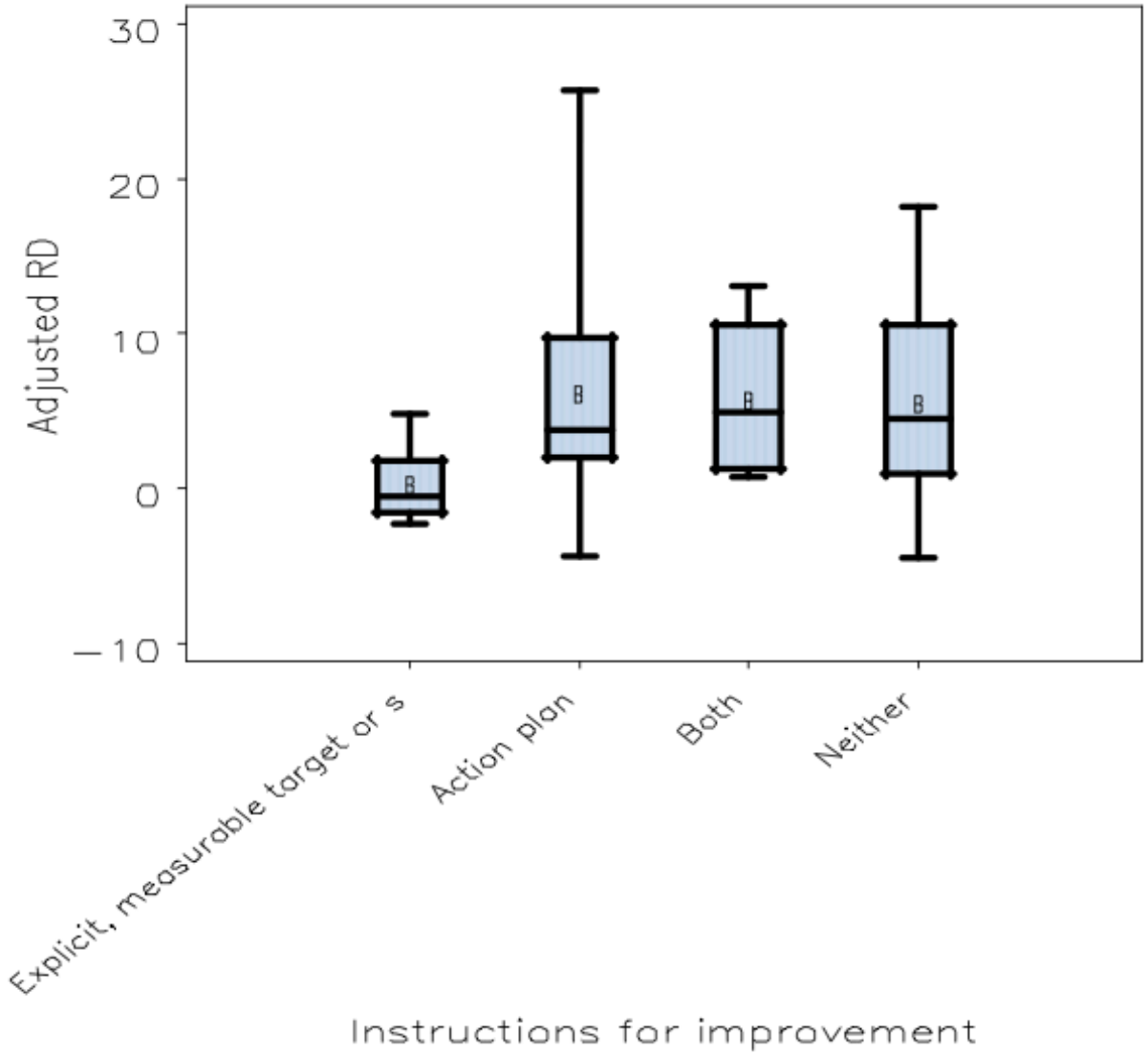
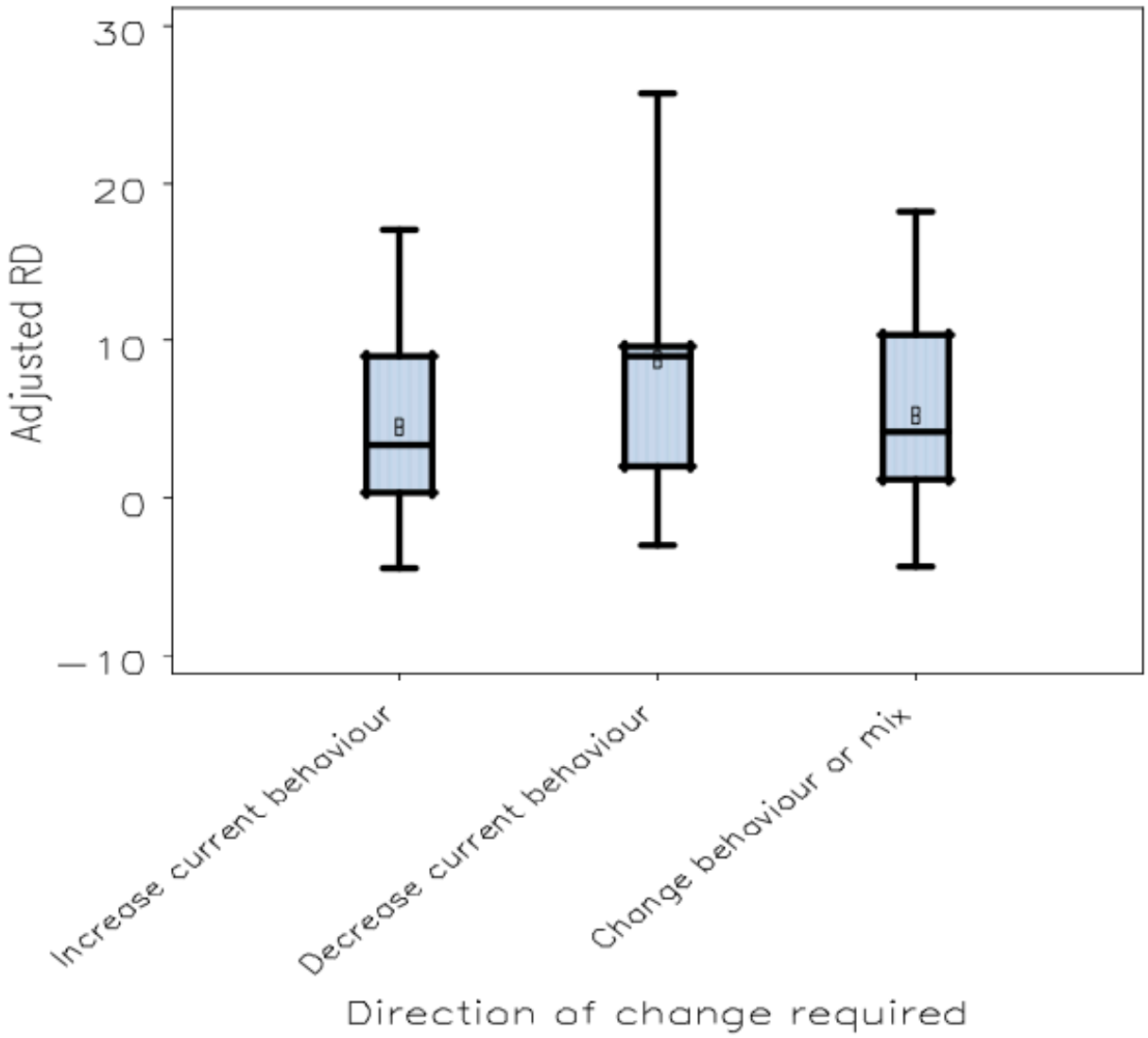
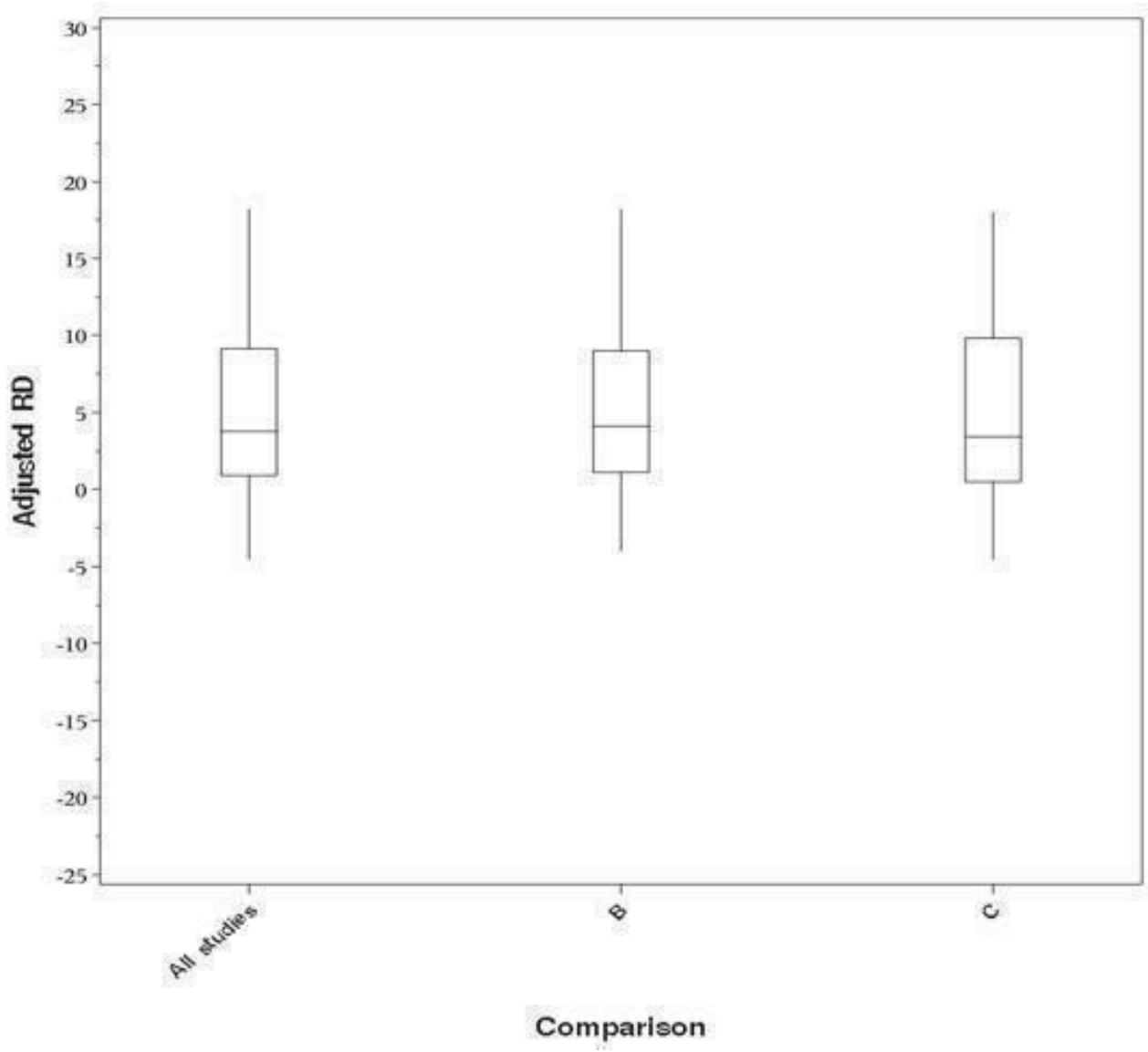


Figure 9. Box plot: comparing adjusted risk difference by direction of change required by the feedback



Although the multifaceted studies appeared to have a larger median effect size, when comparing the mean estimate of effect for audit and feedback alone versus audit and feedback in a multifaceted intervention using a univariate analysis we found that the differences were not statistically significant for dichotomous outcomes (estimated absolute difference in adjusted RD = 3.3%; P = 0.27). The similarity in estimated adjusted RD is illustrated in Figure 10. However, there was a significant difference when examining the studies with continuous outcomes (estimated absolute difference in adjusted change relative to baseline control = 24%; P < 0.0001).

Figure 10. Box plot: comparing adjusted risk difference for Comparison B (audit and feedback alone versus usual care) and Comparison C (multifaceted intervention featuring audit and feedback versus usual care)



The sensitivity analysis adding level of analysis (patient versus provider versus cluster) to the model did not lead to any significant changes in the results. In another sensitivity analysis, when studies with a high risk of bias were included in the model, the findings remained consistent, with two exceptions: format (written versus verbal versus both) no longer had a significant effect, but profession of recipient did, with non-physicians performing better than physicians. It was observed that the model-based estimated effect sizes increased when the high risk of bias studies were included, suggesting caution is needed when interpreting these results. Given that some of the strata within the model were quite small (e.g. only six comparisons from four studies assessed 'both' goals and action plans), such instability is not surprising.

Exploratory analyses

Exploratory analyses were conducted to examine the importance of context and the complexity of the targeted behaviour on the likelihood that audit and feedback will improve professional practice. Although clinical setting (outpatient versus inpatient versus mixed, other or unclear) was marginally statistically significant in the multivariate meta-regression model (P = 0.037), the estimated effects were similar across inpatient and outpatient settings (inpatient estimated RD = 7.7%; outpatient estimated RD = 7.1%; mixed, other or unclear estimated RD = 3.0%).

When 'targeted behaviour' (prescribing versus laboratory or radiology utilisation versus diabetes or cardiovascular disease management versus other) was added to the meta-regression model, it was statistically significant (P < 0.0001), with estimated RD for prescribing (11.1%) larger than diabetes or cardiovascular disease (5.9%), laboratory or radiology testing (4.2%), or other (4.7%). In that model, the 'direction of change required' (increase current behaviour versus decrease versus mix/other) was no longer statistically significant (P = 0.525) and the estimates for some other variables changed (see Table 3). We then conducted

meta-analyses on the subgroups of studies that focused on the targeted behaviours of interest. For prescribing, the weighted median adjusted RD was 13.1% (IQR 3% to 17%) based on 26 comparisons with dichotomous outcomes at unclear or low risk of bias. For laboratory or radiology test utilisation, the weighted median adjusted RD was -0.1% (IQR -0.1% to 6.5%) based on three comparisons, and for trials focusing on the management of diabetes or cardiovascular disease, the weighted median adjusted RD was 0.5% (IQR -0.5% to 3.4%) based on 14 comparisons.

Comparison B. Audit and feedback alone compared to no intervention

A total of 82 comparisons from 65 studies were included in this comparison. Nine comparisons from six trials did not report baseline data and 13 comparisons from 10 trials assessed patient outcomes as a primary outcome, leaving 59 comparisons from 48 studies for the analyses.

For studies with audit and feedback alone targeting professional practice with dichotomous outcomes, there were nine comparisons from seven studies excluded due to high risk of bias, leaving 32 comparisons from 26 studies for the primary analysis. These studies included 759 groups of health providers (from 12 cluster trials) and 1617 health professionals (from 14 trials allocating individual providers). The weighted median adjusted RD was 3.0% (IQR 1.8% to 7.7%). Including the studies at high risk of bias resulted in no change to the estimate of effect.

For studies with audit and feedback alone targeting professional practice with continuous outcomes, there were five comparisons from four studies excluded due to high risk of bias, leaving 14 comparisons from 13 studies for the primary analysis. These studies included 348 groups of health providers (from eight cluster trials) and 494 health professionals (from five trials allocating individual providers). The weighted median adjusted change relative to baseline control was 1.3% (IQR 1.3% to 11.0%). Including the studies at high risk of bias studies in the sensitivity analysis also resulted in a weighted adjusted change relative to baseline control of 1.3% (IQR 1.3% to 20.1%).

Comparison C. Audit and feedback as the core feature of a multifaceted intervention compared to no intervention

A total of 90 comparisons from 65 studies were included in this comparison. Seven comparisons from six trials did not report baseline data and 13 comparisons from nine trials assessed patient outcomes as a primary outcome, leaving 70 comparisons from 50 studies for the analyses.

For studies with multifaceted interventions featuring audit and feedback targeting professional practice with dichotomous outcomes, there were seven comparisons from seven studies excluded due to high risk of bias, leaving 50 comparisons from 32 studies for the primary analysis. These studies included 1574 groups of health providers (from 26 cluster trials) and 480 health professionals (from seven trials allocating individual providers). The weighted median adjusted RD was 5.5% (IQR 0.4% to 16%). Including high risk of bias studies in the sensitivity analysis resulted in a revised weighted adjusted RD = 6.5% (IQR 0.5% to 16%).

For studies with multifaceted interventions featuring audit and feedback targeting professional practice with continuous outcomes, there were 12 comparisons from 11 studies for the primary analysis. These studies included 317 groups of health providers (from seven cluster trials) and 111 health professionals (from four trials allocating individual providers). The weighted median adjusted change relative to baseline control was 26.1% (IQR 12.7% to 26.1%). There were no studies in this group with high risk of bias.

Comparison D. Different ways of providing audit and feedback (head-to-head comparisons)

Seventeen trials included 16 head-to-head comparisons of different ways of providing audit and feedback. For each comparison, we determined the adjusted RD or the adjusted change relative to baseline control. This is reported below in addition to any statistical comparisons conducted by the authors of a particular study (e.g. odds ratios or P values) to provide a standard measure of effect across all comparisons in this review.

Peer comparison

Søndergaard 2002 and Wones 1987 each found small differences when adding peer comparison data to the audit and feedback for asthma management (adjusted RD = 2%) or inpatient laboratory test utilisation (adjusted change relative to baseline control = 5%), respectively. Kiefe 2001 compared audit and feedback featuring a mean score of peers with feedback that featured an “achievable benchmark” (the mean score of the top 10% of peers). They found that the achievable benchmark group improved quality of care for diabetic patients (median adjusted RD = 3%, IQR = 2% to 4%). In particular, statistically significant increases were observed for influenza vaccination (OR 1.54, 95% CI 1.26 to 1.96), foot examination, (OR 1.33, 95% CI 1.05 to 1.69) and haemoglobin A1C measurement (OR 1.33, 95% CI 1.04 to 1.69), while cholesterol measurement (OR 1.20, 95% CI 0.95 to 1.51) and triglyceride measurement (OR 1.15, 95% CI 0.92 to 1.44) had non-statistically significant increases. In contrast, Schneider 2008 found that identifying top performers in feedback presented in a quality circle (i.e. learning collaborative) did not lead to improvements in management of asthma (adjusted RD = -5%, high risk of bias).

Presentation of feedback and inclusion of additional information

Mitchell 2005 found that feedback was slightly more effective for control of blood pressure if it presented information in a way that identified patients at higher risk, suggesting that action for such patients should be prioritised (adjusted RD = 2%; OR 1.72, 95% CI 1.09 to 2.70). (This is a ‘patient’ outcome due to the role of patient-specific factors in achieving control of hypertension. Larger effects on professional practice outcomes might be expected.)

Two studies directly compared including a small amount of extra information to not including that information. Buntinx 1993 added brief advice to typical feedback. They found similar effects for the quality of pap smears (adjusted RD = 1%; no statistical test reported for this comparison). Curran 2008 added ‘Pareto’ and ‘cause and effect' charts’ to help recipients identify barriers and focus improvement efforts. They did not find a statistically significant difference in rates of methicillin-resistant *Staphylococcus aureus* infections in hospital wards (adjusted change = 5%, high risk of bias, patient outcome; P = 0.46).

Two studies tested the type and amount of data used for the feedback reports. Gullion 1988 compared feedback regarding

blood pressure laboratory values, and medications from chart audits to feedback regarding blood pressure and adherence to medication and lifestyle recommendations from patient surveys. They reported no differences in blood pressure control (adjusted RD = 2%, patient outcome; no P value reported for this comparison). Herrin 2006 compared feedback based on administrative data to this plus additional, patient-specific clinical data from medical records. They also did not find a statistically significant difference in the proportion of adequate glucose control (adjusted RD = 1.9%, patient outcome; P = 0.97).

Source and delivery

Four studies directly tested whether feedback should be delivered by mail (written) or in-person (verbally). Rubin 2001 compared written feedback delivered only to the hospital administration with the addition of verbal feedback at staff meetings. They did not find a difference in appropriateness of red blood cell transfusions (adjusted RD = -2%; no statistical test reported for this comparison). Sauaia 2000 found differences that were not statistically significant between verbal feedback in a large group setting by an expert cardiologist and written feedback for improving eight quality of care outcomes related to acute management of myocardial infarction (median adjusted RD = 7%; P value for each outcome > 0.05). Batty 2001 compared similar interventions for in-hospital benzodiazepine prescriptions. The verbal presentation was more effective than the written feedback (adjusted RD = 24%, high risk of bias). Finally, Anderson 1994 found little or no difference when they compared feedback given to large groups as part of a CME (continuing medical education) program, with and without sending individualised feedback reports to participants for prophylaxis of venous thromboembolism (adjusted RD = 0%; no P value reported for this comparison).

Two studies directly tested the effects of who delivered the feedback. Ward 1996 compared audit and feedback delivered by a physician-peer with audit and feedback delivered by a nurse. They found that peer-physician feedback led to non-statistically significant improved management of diabetes (adjusted change relative to baseline control = 12%; P value reported as “NS”). They also noted that the physician interviews were longer (25 minutes versus 14 minutes; P < 0.001) and that there was a significant variation in effect size across the different physicians providing the outreach. Similarly, Van den Hombergh 1999 found that mutual feedback by physician-peers (ie. each physician provides and receives feedback in turn) improved outcomes as measured by 33 indicators of practice management compared with unidirectional feedback by a non-physician (median adjusted RD = 5%; no overall statistical test reported).

Recipient participation

Two studies directly tested the role of recipient participation. Sommers 1984 found that participation in criteria setting prior to the feedback resulted in worse management of anaemia in hospitalised patients (adjusted RD = -21%, high risk of bias; OR = -3.36, P = 0.002). Conversely, Brady 1988 found that when resident physicians conducted a self-audit at baseline, it led to improvements compared with simply receiving the data for mammographic screening rates (adjusted RD = 8%; no OR reported, P value reported as < 0.05) but not to a statistically significant improvement for influenza vaccination rates (adjusted RD = 1.5%; no OR reported, P = 0.17).

Comparison E. Audit and feedback combined with complementary interventions compared to audit and feedback alone

Fifty-three comparisons from 43 trials were included. Below, the results of these comparisons are summarised within categories related to the 'type' of intervention that audit and feedback was combined with when comparing to audit and feedback alone. We acknowledge that some of the multifaceted interventions may fit into multiple categories, but only describe the findings from each trial once. Multi-arm studies may be described in multiple sections corresponding with the type of comparison. Due to the variation in outcome type (dichotomous, continuous, patient, provider) across the studies, we were unable to conduct quantitative meta-analyses, with the exception of trials comparing audit and feedback with educational outreach to audit and feedback alone (see below). For each comparison, we determined the adjusted RD or the adjusted per cent change relative to baseline performance in the audit and feedback alone arm. This is reported below in addition to any statistical comparisons conducted by the authors of a particular study (e.g. odds ratios or P values) to provide a standard measure of effect across all comparisons in this review.

Audit and feedback with reminders compared to audit and feedback alone

Seven studies evaluated adding reminders to audit and feedback. Two of these aimed to reduce outpatient test-ordering. In a 2x2 factorial trial, Eccles 2001 found that adding reminders to audit and feedback reduced x-ray utilisation (adjusted change relative to baseline control = 46%; no P value reported). In another 2x2 factorial trial Thomas 2006 found that feedback and reminders both significantly reduced blood test utilisation and that the effect seemed to be additive, but not synergistic (adjusted change relative to baseline performance in the audit and feedback alone arm -2%; OR = 0.78, 95% CI 0.71 to 0.85 for both versus OR = 0.87, 95% CI 0.81 to 0.94 for reminders alone, no P value reported).

Two studies combined reminders with audit and feedback in an attempt to improve management of diabetes. Phillips 2005 found little or no differences in haemoglobin A1C, systolic blood pressure, and low-density lipoprotein cholesterol levels (median adjusted change relative to baseline performance in the audit and feedback alone arm = 2%; no P value reported). Ziemer 2006 assessed clinical inertia in diabetes and found that the combination of reminders and feedback had a greater effect on treatment intensification than feedback alone (adjusted RD = 7.25%; no P value reported).

Tierney 1986 in a complex factorial trial with active controls found that reminders together with audit and feedback were more effective than feedback alone for provision of preventive services by internal medicine trainees (unadjusted RD = 8.0%; no P value reported). Baker 1997 found improvement in the management of chronic benzodiazepine prescriptions (median adjusted RD = 1.7%, high risk of bias; no overall statistical test conducted). Finally, Boekeloo 1990 found a significant decline in the quality of cholesterol management in hospital when reminders were combined with feedback compared with feedback alone (median adjusted RD = -8%, high risk of bias; no overall statistical test conducted).

One trial, Bahrami 2004, compared audit and feedback with a computer decision support system to audit and feedback alone to improve the management of impacted molars; neither intervention

produced a statistically significant improvement (adjusted RD = 6%; no P value presented for this comparison).

Audit and feedback with educational outreach compared to audit and feedback alone

We found 24 comparisons from 19 studies that compared audit and feedback alone to the combination of audit and feedback and educational outreach (also known as academic detailing). For the 15 studies with dichotomous outcomes focusing on professional practice, the weighted median adjusted RD for audit and feedback with outreach versus feedback alone was a 0.7% increase in desired practice (IQR -1.1% to 5.1%). For the four studies with continuous outcomes, the median adjusted change relative to baseline control was 27% (IQR 0% to 40.5%).

The PINCER trial (Avery 2010) had a median adjusted RD = 1.6 across three outcomes related to safe prescribing. However, in their multivariable model they found that educational outreach by a pharmacist reduced unsafe prescribing practices by GPs compared to feedback alone for the primary outcomes of NSAID (non-steriodal anti-inflammatory drug) use without PPI (proton-pump-inhibitor) (OR 0.58, 95% CI 0.38 to 0.89), beta-blocker use in asthmatics (OR 0.73, 95% CI 0.58 to 0.91), and ACE (angiotensin-converting enzyme) or diuretic use without electrolyte measurements (OR 0.51, 95% CI 0.34 to 0.78). The educational outreach in Moher 2001 focused on showing primary care providers how to utilise the feedback reports to develop and implement systematic patient recall systems. This resulted in an improvement (adjusted RD = 22%; P = 0.002) in the proportion of patients with adequate assessment of cardiovascular risk factors, but no differences in actual treatment. Ward 1996 also found that outreach led to small but statistically significant improvements in diabetes care compared with postal feedback alone (adjusted change for relative to baseline performance in the audit and feedback alone arm = 35%; P < 0.001).

Two 2x2 factorial studies in Sudan both found small effects on inappropriate antibiotic prescribing with academic detailing compared to audit and feedback alone (Awad 2006 - adjusted change relative to baseline performance in the audit and feedback alone arm = 29%, P < 0.001; Eltayeb 2005 - adjusted RD = 9.2%, no P value reported).

Six other studies comparing educational outreach plus feedback to audit and feedback alone had mixed findings. McClellan 2004 found small, but potentially clinically meaningful improvements in the management of dialysis by adding a multifaceted intervention including educational outreach to feedback (difference in mean urea reduction ratio: P = 0.002), but no statistically significant improvement in the primary outcome (proportion of patients with urea reduction ratio > 65%: adjusted RD = 0.70%; P = 0.8). Rask 2001 found that outreach improved diabetes care for only one of six professional outcomes (median adjusted RD = 9.5%, high risk of bias; no overall P value reported) but not in any of three patient outcomes (median adjusted RD = 0%, high risk of bias; no P value reported for this comparison). Siriwardena 2002 found that only two of seven outcomes related to immunisation rates improved (median adjusted RD = 5%, no overall P value reported). Kinsinger 1998 combined audit and feedback with educational outreach aiming to help primary care providers improve office systems to increase breast cancer screening rates. The intervention did improve the office systems and found an increase in the proportion of patients discussing mammograms (adjusted RD = 4.75%; P = 0.01), but not a statistically significant difference in actual mammography rates (P = 0.56) compared with feedback alone. Likewise, Mold 2008 found that academic detailing led to increased implementation of a variety of quality improvement processes in primary care (e.g. standardised protocols), but these efforts translated into a statistically significant improvement in only one of six preventive services measured (median adjusted RD = 8%, no overall P value reported). Finally, Ornstein 2004 found statistically significant improvement in only two of 21 outcomes related to preventive cardiovascular care in the primary care setting and a difference in overall improvement that was not statistically significant (adjusted RD = 5.5%, P > 0.2).

Opinion leaders were explicitly identified to provide the educational outreach in three studies. Soumerai 1998 found improvements in two of four outcomes related to management of acute myocardial infarction (median adjusted RD = 8.5; no overall P value presented). Laskshminarayan 2010 found significant improvement in two of 10 outcomes related to management of acute ischaemic strokes in hospital, but no overall effect (median adjusted RD = 4%; P value reported as non-significant). Guadagnoli 2000 found no differences for breast cancer treatment (adjusted RD = -2%; no P value reported).

The final six studies found no statistically significant effects when adding educational outreach to audit and feedback for the following outcomes: a global quality score incorporating screening, diagnosis, and management in primary care (Borgiel 1999: adjusted RD = -0.2%; no P value reported); prescribing statins and anticoagulants for high cardiovascular risk patients in primary care (Naughton 2007: median adjusted RD = -0.5%; no overall P value reported); antibiotic prescribing in primary care (Naughton 2009: adjusted change relative to baseline performance in the audit and feedback alone arm = 0%; P = 0.33); and management of urinary incontinence by nurses in primary care (Cheater 2006: median adjusted RD = -3.6%; no overall P value reported). Rantz 2001 also found no effect of outreach on nursing home care (median adjusted RD = -1.1%; no overall P value reported), although a subgroup analysis showed that those who actively participated in the outreach did seem to improve.

Audit and feedback plus other educational interventions compared to audit and feedback alone

Four studies tested the combination of small group education with audit and feedback compared to audit and feedback alone. Herbert 2004 compared combining feedback with problem-based learning groups in primary care to feedback alone and found the combination had a greater effect on appropriate use of antihypertensives (adjusted RD = 7.4%; P value not reported). Also in primary care, Verstappen 2004 compared groups that focused on identifying gaps and developing quality improvement plans for decreasing total laboratory tests ordered to feedback alone and found the groups to be more effective (adjusted change relative to baseline performance in the audit and feedback alone arm = 9%, P = 0.005). However, in the hospital setting Kritchevsky 2008 found that adding a quality improvement collaborative to feedback alone did not improve the utilisation of antibiotics within one hour prior to surgery (adjusted RD = -3.6%; absolute risk reduction (ARR) -3.8, 95% CI -13.9 to 6.2). Likewise, Filardo 2009 found that education regarding continuous quality improvement had no statistically significant impact on hospital-based quality

indicators for pneumonia or heart failure compared to feedback alone (median adjusted RD = -1.7%; P = 0.47), although the authors reported that this finding may be due to poor participation in the intervention group.

Hayes 2001 performed a study comparing written feedback with feedback enhanced by the participation of a trained physician, quality improvement tools, and a project liaison for anticoagulant management of venous thrombosis. The multifaceted intervention did not have a statistically significant effect on the quality of care for venous thrombosis (median adjusted RD = 1%; P values > 0.2 for each of five process outcomes). Hayes 2002 conducted a very similar trial targeting heart failure, again finding no statistically significant effect (median adjusted RD = -1%; no P values reported). These studies did not seem to meet strict definitions for either educational outreach or opinion leaders, but did have many similar aspects.

The effect of adding a seminar to audit and feedback was tested in three studies. Both Eltayeb 2005 (adjusted RD = 7.1%; pre-post change for seminar + feedback: 11.6, 95% CI 6.6 to 16.7 versus pre-post change for feedback alone: 3.8, 95% CI -1.2 to 8.8) and Awad 2006 (adjusted change relative to baseline control = 26%; P < 0.001) found that adding seminars to audit and feedback reduced inappropriate prescribing of antibiotics in Sudan. Robling 2002 found minimal difference in compliance with guidelines for MRI (Magnetic Resonance Imaging) of the lumbar spine or knee (unadjusted RD = 4%, high risk of bias; no P value reported).

Finally, three studies tested written educational materials. Everett 1983 found that combining written education regarding costs with feedback regarding laboratory use seemed to decreased test utilisation compared to audit and feedback alone (unadjusted difference = 22.3%, high risk of bias; no P value reported). Marton 1985 also found that offering a manual outlining laboratory costs reduced laboratory test utilisation compared to feedback alone (adjusted change relative to baseline performance in the audit and feedback alone arm = 33%; no P value reported). Conversely, Hershey 1988 found no significant effect on prescription rates when attaching to feedback a newsletter outlining advantages, disadvantages, and indications of treatment options (adjusted change relative to baseline performance in the audit and feedback alone arm = 8%; no P value reported).

Audit and feedback with case management or organizational interventions compared to audit and feedback alone

Four trials compared audit and feedback with team changes or case management-type interventions to audit and feedback alone. Moher 2001 compared mailed feedback to feedback plus a nurse recall system in a three-arm study. The nurse recall system improved the proportion of patients with adequate assessment of cardiovascular risk factors compared to feedback alone (adjusted RD = 33%; ARR = 33, 95% CI 19 to 46). However, this difference was not reflected in clinical outcomes, such as blood pressure or cholesterol. Similarly, Herrin 2006 found that adding a diabetes resource nurse resulted in minimal changes in glucose control when compared to two different types of feedback alone (adjusted RD = 3.1%, 1.2%; all comparisons reported with P values > 0.1). Using a more intensive intervention, Svetkey 2009 tested the addition of chronic disease group visits and case management to a feedback intervention but found little or no additional effect at 18 months for mean systolic blood pressure (adjusted change relative to baseline performance in the audit and feedback alone arm = 1%; no P values reported).

One study added a telephone follow-up to audit and feedback targeting pneumococcal vaccine coverage (Quinley 2004). This was an administrative task that encouraged use of the feedback reports and required no clinical expertise and the intervention resulted in little or no difference in vaccine use across the two subgroups of physicians analysed (median adjusted RD = 0.97%; P values 0.07, 0.09).

Audit and feedback with financial incentives compared to audit and feedback alone

Two studies compared audit and feedback to audit and feedback plus incentives. Fairbrother 1999 had three arms that compared audit and feedback alone to audit and feedback plus an one-off "financial bonus" based on up-to-date coverage for four immunisations, and audit and feedback plus "enhanced fee for service" (five dollars for each vaccine administered within 30 days of its due date). Rates of immunisation improved from 29% to 54% coverage in the bonus group after eight months (adjusted RD: 12.7%; no P value comparing bonus group to feedback alone). The enhanced fee-for-service group decreased performance relative to feedback alone (adjusted RD -8.3%; no P value for this comparison). A separate study (Hillman 1999) found that adding incentives to audit and feedback did not improve the implementation of paediatric preventative care guidelines (adjusted RD -5.4%, no P value reported).

Audit and feedback with patient-mediated interventions compared to audit and feedback alone

Five trials compared audit and feedback plus patient educational materials with audit and feedback alone and only one showed a positive effect in favour of adding patient education to audit and feedback. Mainous 2000 was a four-arm study that found adding patient educational pamphlets to audit and feedback had little or no influence on antibiotic prescribing for respiratory infections (adjusted RD = 0%; no P value reported for this comparison). Similarly, Schectman 2003 found that patient pamphlets and videos did not improve management of low back pain compared with feedback alone, probably because it was poorly adopted (raw data not reported, patient intervention described as not effective). Buffington 1991 found that mailed patient reminders resulted in little or no difference from weekly feedback alone for influenza vaccination rates (adjusted RD = 1%, high risk of bias; P value reported as 'not significant'). O'Connor 2009 found that mailed information with reminders to patients with diabetes did not increase the effectiveness of a feedback intervention for control of haemoglobin A1C (adjusted change relative to baseline performance in the audit and feedback alone arm = -1%; no P value for this comparison). Weitzman 2009 found that the addition of patient reminders to feedback using both a letter and phone-call to urge comprehensive follow-up resulted in improved control of diabetes based on achieving glucose, cholesterol and blood pressure targets (median adjusted RD = 4.4%; OR = 2.4, P < 0.01).

Comparison F. Other interventions compared to audit and feedback

Twenty two comparisons from 20 trials were included in this comparison. Below, the results of these comparisons are summarised within categories related to the 'type' of intervention

that audit and feedback was combined with when comparing to audit and feedback alone. We acknowledge that some of the multifaceted interventions may fit into multiple categories, but only describe the findings from each trial once. Multi-arm studies may be described in multiple sections corresponding with the type of comparison. Due to the variation in outcome type (dichotomous, continuous, patient, provider) across the studies, we were unable to conduct quantitative meta-analyses. For each comparison, we determined the adjusted RD or the adjusted percent change relative to baseline performance in the audit and feedback arm. This is reported below in addition to any statistical comparisons conducted by the authors of a particular study (e.g. odds ratios or P values) to provide a standard measure of effect across all comparisons in this review.

Reminders compared to audit and feedback

Audit and feedback was compared to reminders in eight studies. Eccles 2001 found that educational reminders appended to radiology reports were more effective than twice yearly feedback to general practitioners for reducing overall radiology requests (median adjusted change relative to baseline performance in audit and feedback arm 15%; pre-post difference in rate for reminders = 1.57, 95% CI 0.6 to 2.5 and pre-post difference for feedback = 0, no P value for this comparison). Tierney 1986 also found that reminders were superior to monthly feedback to medical residents for improving delivery of a variety of preventive services (unadjusted RD 4.5%, no P value reported).

In Thomas 2006, feedback led to greater reductions in the number of laboratory tests ordered compared with reminders although the model-based analyses suggested similar effects (adjusted change relative to baseline performance in audit and feedback arm = 12%; OR for feedback = 0.87, 95% CI 0.81 to 0.94, OR for reminders = 0.89, 95% CI 0.83 to 0.93). In Ziemer 2006, feedback was more effective than reminders for reducing clinical inertia in diabetes, measured as the proportion of visits with action taken to improve glucose control (adjusted RD = 6%; P < 0.01). Finally, Boekeloo 1990 found that audit and feedback was superior to reminders for inpatient cholesterol management (median adjusted RD = 15%, high risk of bias; no P value for this comparison). Grady 1997 found little or no difference between the interventions in rate of mammography referral (adjusted RD = -1%; P value reported as not significant) and Phillips 2005 found minimal difference in management of diabetes (adjusted RD = -0.1%; no P value for this comparison).

Bahrami 2004, compared audit and feedback to a computer decision support system to improve the management of impacted molars. Neither intervention was shown to be effective (adjusted RD = 2%; no P value reported for this comparison).

Educational outreach compared to audit and feedback

Lomas 1991 compared audit and feedback to the use of local opinion leaders to implement guidelines for the management of women with a previous caesarean section in a high quality study. The opinion leader group increased the proportion of women offered a trial of labour and the audit and feedback group did not (unadjusted RD = 17.9%; P = 0.002). Cheater 2006 found somewhat favourable effect for audit and feedback compared to the educational outreach arm, but their models revealed no evidence for either arm in the management of urinary incontinence by nurses situated in family practices and intervention (median adjusted RD = -3.9%; ARR = -2.3%, 95% CI -6.3 to 1.7 for feedback versus ARR = 0.9%, 95% CI -3.3 to 5.1 for outreach).

Other educational interventions compared to audit and feedback

Two studies directly compared seminars to audit and feedback. Robling 2002 did not find a statistically significant difference between feedback and a seminar in appropriateness of MRI requests of the lumbar spine or knee, (unadjusted RD = 12%, high risk of bias; concordance = 67%, 95% CI 52 to 81% for feedback versus 79%, 95% CI 66 to 92% for seminar, no P value reported). Holm 1990 found that a seminar was more effective than audit and feedback for reducing benzodiazepine prescriptions (adjusted change relative to baseline performance in the audit and feedback arm = 22%; P = 0.03).

Herbert 2004 found that practice-based small group learning similarly effective as postal audit and feedback amongst family physicians for increasing appropriate use of antihypertensives (adjusted RD = 0.8%; no P value reported for this comparison).

Finally, two studies tested printed educational materials. Everett 1983 found that printed materials regarding costs of laboratory tests did not lead to changes in laboratory test utilisation, but audit and feedback actually increased utilisation (unadjusted RD = -12.9%, high risk of bias; no P value reported). However, Marton 1985 found that neither a manual outlining costs nor feedback every two weeks on laboratory expenditures significantly reduced laboratory test utilisation (adjusted change relative to baseline performance in the audit and feedback arm = 6%; P value reported as non-significant).

Case management or organizational interventions compared to audit and feedback

When Svetkey 2009 compared chronic disease group visits and case management to audit and feedback, no effect was found for either intervention on systolic blood pressure at 18 months (adjusted change relative to baseline performance in the audit and feedback arm = -1%; effect for feedback = 0.3 mm Hg, P = 0.81 and effect for case management = -0.2 mm Hg, P = 0.89). Claes 2005 did not find a difference between feedback and either point-of-care testing or rapid clinical decision support from the laboratory for keeping patients within target INR (International Normalized Ratio) for their oral anticoagulation, although all interventions seemed to be effective (adjusted change relative to baseline performance in the audit and feedback arm = 4% for both comparisons; P = 0.13 for difference across all arms).

Financial Incentives compared to audit and feedback

Martin 1980 compared incentives to audit and feedback to reduce test-ordering in hospitals. Incentives were less effective than audit and feedback at reducing test-ordering (adjusted change relative to baseline performance in the audit and feedback arm = -41%; P value reported as < 0.05).

Patient-mediated interventions compared to audit and feedback

Three studies directly compared patient-mediated interventions with provider-directed audit and feedback; none found a statistically significant difference in outcomes. Mainous 2000 compared patient educational pamphlets to feedback, finding little or no difference between groups in antibiotic prescribing rates

(adjusted RD = 2%, no P value reported for this comparison). Schectman 2003 did not find a statistically significant effect of patient pamphlets and videos on the management of low back pain. The details of the results of this group compared to the feedback group were not reported. Finally, one study (O'Connor 2009) compared a patient intervention featuring a postal letter to each patient summarising their diabetes-related risk factors and offering suggestions for improvement with a physician intervention featuring audit and feedback plus reminders. No improvement in haemoglobin A1C level was found (adjusted change relative to baseline performance in the audit and feedback arm = -1%; no P value for this comparison).

DISCUSSION

Summary of main results

Audit and feedback can be a useful intervention to improve health professionals' compliance with desired practice. The median adjusted risk difference (RD) of compliance with desired practice was a 4.3% absolute increase in desired practice (IQR 0.5% to 16%) when considering any trial in which audit and feedback was considered the core, essential aspect of the intervention, compared to no audit and feedback. For continuous variables, we found that the weighted median adjusted change relative to the performance of the control group at baseline was a 1.3% increase in compliance with desired practice (IQR 1.3% to 23.2%). Although the median effect may be perceived as relatively small, the 75th percentile effect size is much larger (16% absolute improvement in health professionals compliance with desired behaviour), suggesting that audit and feedback, when optimally-designed and used in the right context, can play an important role in improving professional practice.

There are a number of plausible explanations why some interventions were more effective than others and we tested some of the hypothesised variables in a meta-regression. As in the previous versions of this review, we found that baseline performance was associated (inversely) with the effectiveness of audit and feedback. The meta-regression provides indirect evidence that five feedback characteristics are also associated with the effectiveness of audit and feedback interventions. Specifically, our findings indicate that feedback will be most effective when provided from a source that is a 'supervisor or senior colleague', and delivered at least 'monthly', in both a 'verbal and written' format, aiming to decrease rather than increase provider behaviours, and offers instructions with 'both explicit goals and a specific action plan'. However, the ability to make firm conclusions from the analysis of heterogeneity is hindered both by the indirect nature of the comparisons and by the non-specific nature of the components of those variables. For instance, while it appears that verbal feedback is the least effective format, such 'verbal' feedback could have been a lecture to a large group or a one-to-one discussion. Likewise, while it appears that a 'supervisor or colleague' is the most effective source, this finding may depend on whether or not the colleague is a respected opinion leader. In addition, the difference in effect between interventions aiming to decrease or increase behaviours vanished when the targeted behaviour was analysed in the exploratory analysis. Therefore, the results of our meta-regression should be interpreted cautiously.

Seventeen studies provided direct, randomised comparisons of different ways of providing audit and feedback; only four of these trials were published after 2003. Based on these comparisons and also based on indirect comparisons across studies it is difficult to determine what other features of audit and feedback have an important impact on its effectiveness. For example, we found conflicting evidence regarding the role of peer comparisons. Kiefe 2001 indicated that comparing to the top 10% of peers might be an improvement over comparing to the mean, but Schneider 2008 found that identifying top performers in the context of a quality circle did not increase the effectiveness of feedback. The difference may reflect the role of explicit goal/target setting in determining the reaction to feedback (Locke 2002; Carver 1982). Active participation in goal-setting may also play an important role (BMJ 1992). Although there are theoretical reasons why some forms of audit and feedback might be more effective than others, there remains a need to operationalise and directly compare different approaches to improving the design and delivery of audit and feedback. For now, decisions about when to provide audit and feedback must largely be guided by pragmatic considerations and hypotheses based on a priori theory.

In addition to the design of the intervention itself, it is likely that the characteristics of the context and the recipients might influence the effectiveness of feedback. Furthermore, feedback might also be best suited for changing specific types of behaviours; for example, more complex targeted behaviours might be harder to change by providing feedback. When we attempted in the previous version of this review to include the complexity of the targeted behaviour as a variable in our meta-regression, we did not find a statistically significant association between the complexity of the targeted behaviour and the effectiveness of feedback, possibly because it was difficult to reliably assess complexity.

In this review, we conducted an exploratory analysis for a small number of targeted behaviours (prescribing, test-ordering, and management of diabetes or cardiovascular disease) chosen because they were frequently targeted in feedback trials. We found a relatively large effect for prescribing (median adjusted RD 13.1%) compared with test-ordering (-0.1%) and management of diabetes or cardiovascular disease (0.5%). A plausible explanation for this difference is that prescribing is typically not a complex behaviour and may be perceived as important, whereas test-ordering may be perceived as less important (and might be more complex) and disease management is typically more complex. However, within the diabetes and cardiovascular disease subgroup there was great variation in the targeted behaviours. This is also true of the prescribing and the test-ordering subgroups. In some trials, the intention was to increase prescribing, test-ordering or referrals (addressing under-use), while in others the goal was to reduce utilisation (addressing over-use). It is important that future trials consider carefully the intended target of the intervention and precisely describe the targeted behaviours, ideally including an assessment of their complexity and perceived importance. Although our analysis suggests that audit and feedback might be highly effective for improving prescribing (and less effective for test-ordering or disease management), this was an exploratory analysis and there remains a great deal of uncertainty regarding which clinical or behavioural targets would be most appropriate for audit and feedback.

The previous version of this review investigated the impact of audit and feedback when used as part of a multifaceted intervention, finding little evidence of enhanced effectiveness, consistent with

other reviews that have concluded that multifaceted interventions are not necessarily more effective than single strategies (Forsetlund 2009; Grimshaw 2004; O'Brien 2008). In this review, we found that when audit and feedback was combined with other interventions, the effect size of the intervention was larger than when audit and feedback was used alone. This difference was statistically significant for studies with continuous outcomes but not with dichotomous outcomes. The results were also inconsistent with respect to suggesting which combinations of interventions might be most effective. Thus, the added costs of multifaceted interventions need to be weighed against the uncertainty of whether a multifaceted intervention is likely to produce a greater effect. When and how to best combine feedback with other interventions warrants systematic investigation, ideally through a series of comparative trials.

Overall completeness and applicability of evidence

Although the variation in effect size is noteworthy and requires further study, the consistency of median effect size found in this review compared with the previous review, despite changes in methodology, is of interest. While the best way to design and deliver feedback remains uncertain, this review provides greater certainty about its likely effect compared with usual care across a variety of clinical situations. Given the large number of RCTs included in this review and the stability in effect size observed over time, we believe it is unlikely that missing or new trials of audit and feedback versus usual care would substantially alter the estimated median effect of audit and feedback on professional practice. Thus, future trials should aim to determine the best way to deliver audit and feedback in head-to-head trials rather than comparing audit and feedback to usual care.

Quality of the evidence

In most of the included studies, the method of allocation was not clearly indicated in the published report. Although lack of allocation concealment can result in overestimates of effect (Odgard-Jensen 2011), the importance of this criterion in trials where a group of healthcare professionals is randomised at one point in time is not established. In this review, we have given cluster-randomised trials the benefit of the doubt and assumed that there was adequate concealment of allocation for these studies.

Nonetheless, we judged only 32% of the included studies to have a low risk of bias. This compares favourably to the previous version of this review in which only 20% of the included trials had a low risk of bias. However, we judged 18% of the studies included in this review to have a high risk of bias, while in the previous review only 12% were deemed high risk of bias. The lower proportion of studies with unclear risk of bias may represent improved reporting over time. As with the previous review, we found no association between overall risk of bias (low versus unclear) and the estimate of effect.

Potential biases in the review process

In this review, our inclusion criteria required that at least one arm of the trial use audit and feedback as the core, essential feature of the intervention. This was necessary to avoid including trials of multifaceted interventions where feedback was included but where the main effects of the intervention were unlikely to be due to feedback. If some effective multifaceted interventions were inappropriately excluded, this would create a conservative bias (and vice-versa). Although application of this criterion depended on judgements made by the review authors, only eight disagreements occurred between independent reviewers of 282 full-text manuscripts reviewed and all were resolved easily through discussion. Furthermore, the similarity in the estimate of effectiveness for multifaceted interventions featuring audit and feedback between this review (adjusted RD 5.5%) and the previous review (adjusted RD 5.7%) supports the notion that the operationalisation of this criterion did not substantively bias the results.

In earlier reviews of this topic, we considered printed educational materials to have little or no effect on changing professional practice based on information available at the time (Freemantle 1997; Grimshaw 2001). However, recent reviews (Farmer 2008; Grimshaw 2004) found that printed educational materials have a small (but potentially important) effect. By abstracting printed materials as usual care for many studies, we may have created a conservative bias for studies comparing feedback to printed materials, but an overestimation of the effect attributed to audit and feedback in studies where feedback plus printed materials are compared to no intervention. In most studies educational materials were distributed to all groups, thus meeting a pragmatic definition of usual care.

One possible reason for our finding that few studies featured patient outcomes is that we only abstracted primary outcomes and many studies provided patient outcomes as secondary outcomes. This assessment would have been easier to make if more studies clearly stated their primary outcome in general and if more studies had planned to have a patient level outcome as the primary outcome. However, since most studies reporting patient outcomes as secondary outcomes are likely to be under powered to detect a difference in patient outcomes, this is unlikely to have affected our finding that improvements in patient outcomes were at best small. The reason for this is that impacts on patient outcomes depend on the combined effectiveness of feedback on professional practice and the effectiveness of the clinical intervention (delivered as a result of the changed in professional practice). Since the effectiveness of feedback is typically small or moderate (e.g. a 4.3% absolute improvement) and the effectiveness of targeted clinical interventions (changes in practice) is typically moderate, feedback can only be expected to have a small effect on patient outcomes in most circumstances. Thus large trials are needed to reliably measure the impacts of feedback on patient outcomes.

As illustrated in Appendix 1, there are many possible factors related to feedback design that could potentially predict effectiveness. It is certainly possible that we neglected to abstract some important design factors, especially organisational and contextual characteristics. We limited the exploration of such factors for pragmatic reasons (based on feasibility of abstraction) and to limit risk of spurious findings.

We chose to focus on comparisons where it was possible to calculate an adjusted risk difference and adjusted change relative to the baseline control. The adjustments were based on pre-intervention measurements of the outcome in the audit and feedback group. We excluded from the quantitative analyses studies without baseline data because of previous evidence that baseline performance is associated with effectiveness of audit and feedback. Since many studies included small numbers of healthcare professionals, baseline differences were common and unadjusted estimates of effect often differed from the adjusted

estimates. This being said, we acknowledge that, ideally, across a systematic review these baseline differences should cancel each other out (as each imbalance is random); thus the post intervention comparison should be just as useful as the adjusted estimates, as long as studies lacking such data were not systematically different in other respects. Therefore, our choice to exclude studies without baseline measurements from analyses may be regarded as an additional potential limitation of the review.

We weighted the analyses by the number of health professionals involved in each trial. Trials that did not report the number of health professionals involved were weighted by the number of practices/hospitals/communities involved in the trial; this typically occurred when the unit of allocation was a cluster of providers (e.g. practice, hospital, or community) rather than a single provider. This approach may have led to some larger studies with many participants but relatively few clusters being assigned a weight that did not reflect the actual size of the trial.

Agreements and disagreements with other studies or reviews

The previous update of this Cochrane review found similar estimates of effect for audit and feedback on professional practices. It also found that greater "intensity" of feedback was associated with greater effect. However, the assessment of "intensity" simultaneously captured numerous variables and was, therefore, difficult to operationalise as it could not discern which components were most important. In this review, we tested five specific characteristics of feedback design in a meta-regression in an attempt to identify important active ingredients of audit and feedback.

The sources of feedback associated with the lowest effect size were 'professionals standards review organisation' and 'representative of the employer or purchaser'. This fits well with previous qualitative work comparing high and low performing hospitals finding that feedback with a punitive tone seems to be less effective (Hysong 2006). Also of note was the stability in effect across clinical setting and profession of recipient, although the latter finding may be due to the paucity of trials with interventions directed to non-physicians. Our finding that risk of bias was not associated with effect size is consistent with the previous version of this review. In both cases, this may be explained by suboptimal reporting, resulting in many risk of bias domains judged to be 'unclear'.

The findings of this review regarding format and source were consistent with a re-analysis of a previous version of the review (Hysong 2009) and should also be considered in light of the Feedback Intervention Theory (Kluger 1996), which suggests that feedback that directs attention towards acceptable and familiar tasks (as opposed to those that generate emotional responses or cause deep self-reflection) seem most likely to lead to improvement. Our results regarding action-planning are also consistent with the re-analysis of the previous version of the review informed by the Feedback Intervention Theory (Hysong 2009). However, a separate re-analysis of the previous version of this review aiming to test the hypothesis regarding goal-setting and action-planning found too few studies to reach any conclusions (Gardner 2010). Although we hypothesised based on Carlsen 2007 that feedback aiming to increase provider behaviours would be more effective than feedback aiming to decrease behaviours we found that the opposite was true. This suggests that stated preferences with respect to quality improvement interventions should be empirically tested.

In this review, we also found evidence that the targeted behaviour may be associated with the effectiveness of feedback. In particular, we found that feedback aiming to change prescribing habits may be more effective that feedback aiming to improve chronic disease management. A recent review of audit and feedback given to general practitioners regarding diabetes management (Guldberg 2009) included 10 studies with great heterogeneity in outcomes. The authors were unable to conclude which diabetes process measures should be targeted by future interventions and more work is clearly needed in this area.

Previous reviews have looked at factors associated with the effectiveness of audit and feedback and we recognised from the outset that there are far more plausible factors that could alter the effectiveness of audit and feedback than we could test in this review.

Mugford and colleagues (Mugford 1991) identified 36 published studies of information feedback which they defined as the use of comparative information from statistical systems. These authors distinguished passive from active feedback where passive feedback was the provision of unsolicited information and active feedback engaged the interest of the clinician. They also assessed the impact of the recipient of the information, the format of the information and the timing of the feedback. Studies were included if their design used either a historical or a concurrent control group for comparison. The authors concluded that information feedback was most likely to influence clinical practice if the information was presented close to the time of decision-making and the clinicians had previously agreed to review their practice. The results of this review do not support or refute these conclusions. Axt-Adam and colleagues (Axt-Adam 1993) reviewed 67 published papers of interventions (26 studies of feedback) designed to influence the ordering of diagnostic laboratory tests. They reported factors that could be important included the message, the provider of the feedback, the addressee, the timeliness and the vehicle. They concluded that there was considerable variation among different studies and that this variation could be explained in part by the extent, the timing, the frequency, and the availability of comparative information related to peers. They also felt that the practice setting was an important factor. Our findings support many of these conclusions. Buntinx and colleagues (Buntinx 1993) conducted a systematic review of 26 studies of feedback and reminders to improve diagnostic and preventive care practices in primary care. They categorised the information provision that occurred after or during the target performance as feedback whereas information provision that occurred before the target performance was called reminders. Ten of the 26 studies used randomised designs but the quality of the included trials was not reported. The authors concluded that both feedback and reminders might reduce the use of diagnostic tests and improve the delivery of preventive care services. However, they also reported that it was not clear how feedback or reminders work, especially the use of peer group comparisons. Balas and colleagues (Balas 1996) reviewed the effectiveness of peer-comparison feedback profiles in changing practice patterns. They located 12 eligible trials and concluded that profiling had a statistically significant but minimally important effect.

AUTHORS' CONCLUSIONS

Implications for practice

Audit and feedback can be effective in improving professional practice. The effects are generally small to moderate and vary based on the way the intervention is designed and delivered. As with any quality improvement strategy, efforts to change provider practice should be targeted at behaviours for which there is evidence between processes and patient outcomes.

The results of this review suggest that feedback may be more effective when baseline performance is low, when the source is a supervisor or senior colleague, when it is provided more than once, when it is provided both verbally and written, and when it includes both measurable targets and an action plan. In addition, the effect size varies based on the clinical behaviour targeted by the intervention. Although the quality of evidence for these findings is low, it is sensible to provide measurable targets and an action plan when this is practical, since this is unlikely to entail additional costs or harms. On the other hand, pragmatic consideration needs to be given to additional costs associated with providing feedback more frequently, providing both verbal and written feedback, and using a supervisor or colleague to provide feedback, since these features may entail additional costs while the benefit is not certain. The finding related to decreasing provider behaviours may suggest that feedback could be useful in situations where there is a desire to curb over-utilisation, keeping in mind that the source of the feedback should preferably be a senior colleague rather than the payor.

Audit is commonly used to improve accountability, either in the context of governance or as a feature of ongoing quality improvement efforts. The findings of this review suggest that it may be possible to increase the effect of feeding back the results of such audits on professional practice through careful attention to the way the feedback is designed and delivered. Those planning new interventions aiming to change practice should consider audit and feedback alongside other interventions and weigh the potential benefits against the potential challenges with respect to cost and/or logistics.

Implications for research

There are two main research audiences for this review: those who wish to implement and rigorously evaluate the effectiveness of a local audit and feedback intervention and those who wish to examine the underlying cognitions and behavioural control mechanisms that may explain how to best design and deliver these interventions. Like other reviews of quality improvement interventions, we have found limited progress over time in the knowledge of when and how to best conduct audit and feedback interventions (Flodgren 2011; Forsetlund 2009; O'Brien 2008). This suggests an opportunity for improved collaboration between the 'applied' scientists aiming to improve local quality of care and 'basic' scientists aiming to produce generalisable knowledge. In particular, each new audit and feedback intervention may provide an opportunity to incorporate evaluations of different ways of designing and/or delivering the feedback to explore how to optimise this intervention in routine practice settings. To build upon the current evidence base, the field would benefit from more attention to four areas: improved reporting and methods; explicit use of theory, empirical evidence, and logic to develop hypotheses and to design the intervention and comparison arms; a focus on professional practices for which there is compelling evidence of patient benefits with clearly defined primary outcomes; and more head-to-head trials (e.g. comparing different ways of providing feedback).

At a minimum, to contribute to the literature, trials need to be well-designed and clearly reported (Simera 2010). Better reporting of study methods, targeted behaviours, characteristics of participants, and the context are needed (Davidoff 2009). A clear, thorough description of the intervention, ideally with illustrative examples would be useful. Primary outcomes should be important and clearly specified. The results should be adjusted for baseline differences, which are common in cluster-randomised trials, and the analysis should take account of the unit of allocation. Furthermore, trials need to be large enough to detect small effects (especially for changes in patient outcomes), when those effects are considered important.

The field would likely benefit if investigators explicitly built upon knowledge generated from prior trials, systematic reviews, and relevant theory to design audit and feedback interventions. In addition to some of the psychology literature referred to in the background section, the education and the organisational/management literature suggest how the design and delivery of feedback might be optimised to improve performance (see, for example, Shute 2008). Well-designed, mixed methods process evaluations embedded within trials can be useful to explore and provide insights into the complex dynamics underlying the variable effectiveness of audit and feedback. In particular, researchers should examine hypotheses regarding how their audit and feedback intervention will be acted upon in practice.

Finally, although there have been more trials over time directly comparing different ways of conducting feedback interventions, there is a continued need to emphasise this type of head-to-head trial. The cumulated evidence suggests that further two-arm trials comparing feedback with usual care are likely to be of limited value. The focus should shift from whether audit and feedback works better than usual care to discerning ways to optimise the effectiveness of audit and feedback interventions for particular contexts or clinical practices. The utility of future updates of this review will depend on the availability of new, well-designed (and well-reported) trials and on our ability to recognise, abstract, and analyse important explanatory factors.

ACKNOWLEDGEMENTS

We would like to thank Kjetil Olsen for his clerical assistance and all authours who made contributions to earlier versions of this review.

REFERENCES

References to studies included in this review

Anderson 1994 *{published data only}*

Anderson FA Jr, Wheeler HB, Goldberg RJ, Hosmer DW, Forcier A, Patwardhan NA. Changing clinical practice. Prospective study of the impact of continuing medical education and quality assurance programs on use of prophylaxis for venous thromboembolism. *Archives of Internal Medicine* 1994;**154**:669-77.

Avery 2010 *{published data only}*

Avery AJ, Rodgers S, Cantrill JA, Armstrong S, Boyd M, Cresswell K, et al. PINCER trial: a cluster randomised trial comparing the effectiveness and cost-effectiveness of a pharmacist-led IT-based intervention with simple feedback in reducing rates of clinically important errors in medicines management in general practices. A report for the Department of Health Patient Safety Research Portfolio. 2010.

Cantrill J. International Journal of Pharmacy Practice. 2010 Conference:5.

Awad 2006 *{published data only}*

Awad AI, Eltayeb IB, Baraka OZ, Cano-Garcinuno A, az-Vazquez C, Carvajal-Uruena I, et al. Changing antibiotics prescribing practices in health centers of Khartoum State, Sudan; Group education on asthma for children and caregivers: a randomized, controlled trial addressing effects on morbidity and quality of life. *European Journal of Clinical Pharmacology* 2006;**62**:135-42.

Bahrami 2004 *{published data only}*

Bahrami M, Deery C, Clarkson JE, Pitts NB, Johnston M, Ricketts I, et al. Effectiveness of strategies to disseminate and implement clinical guidelines for the management of impacted and unerupted third molars in primary dental care, a cluster randomised controlled trial. *British Dental Journal* 2004;**197**(11):691-6.

Baker 1997 *{published data only}*

Baker R, Farooqui A, Tait C, Walsh S. Randomised controlled trial of reminders to enhance the impact of audit in general practice on management of patients who use benzodiazepines. *Quality in Health Care* 1997;**6**:14-8.

Baker 2003 *{published data only}*

Baker R, Falconer J, Lambert PC. Randomized controlled trial of the effectiveness of feedback in improving test ordering in general practice. *Scandinavian Journal of Primary Health Care* 2003;**21**:219-23.

Baker 2003A *{published data only}*

Baker R, Fraser RC, Stone M, Lambert P, Stevenson K, Shiels C. Randomised controlled trial of the impact of guidelines, prioritised review criteria and feedback on implementation of recommendations for angina and asthma. *British Journal of General Practice* 2003;**53**:284-91.

Balas 1998 *{published data only}*

Balas E, Boren SA, Hicks LL, Chonko AM, Stephenson K. Effect of linking practice data to published evidence: A randomized controlled trial of clinical direct reports. *Medical Care* 1998;**36**:79-87.

Batty 2001 *{published data only}*

Batty G, Oborne CA, Hooper R, Jackson S. Investigation of intervention strategies to increase the appropriate use of antithrombotics in elderly hospital inpatients with atrial fibrillation. *Journal of Clinical Governance* 2001;**9**:115-22.

Beck 2005 *{published data only}*

Beck CA, Richard H, Tu JV, Pilote L. Administrative Data Feedback for Effective Cardiac Treatment: AFFECT, a cluster randomized trial. *JAMA* 2005;**294**:309-17.

Bentz 2007 *{published data only}*

Bentz CJ, Bayley KB, Bonin KE, Fleming L, Hollis JF, Hunt JS, et al. Provider feedback to improve 5A's tobacco cessation in primary care: a cluster randomized clinical trial. *Nicotine and tobacco research* 2007;**9**(3):341-9.

Berman 1998 *{published data only}*

Berman MF, Simon AE. The effect of a drug and supply cost feedback system on the use of intraoperative resources by anesthesiologists. *Anesthesia and Analgesia* 1998;**86**:510-5.

Blais 2008 *{published data only}*

Blais R, Laurier C, Paré M. Effect of feedback letters to physicians and pharmacists on the appropriate use of medication in the treatment of asthma. *Journal of Asthma* 2008;**45**(3):227-31.

Boekeloo 1990 *{published data only}*

Boekeloo BO, Becker DM, Levine DM, Belitsos PC, Pearson TA. Strategies for increasing house staff management of cholesterol with inpatients. *American Journal of Preventive Medicine* 1990;**6**(Suppl 2):51-9.

Bonevski 1999 *{published data only}*

Bonevski B, Sanson-Fisher RW, Campbell E, Carruthers A, Reid ALA, Ireland M. Randomized controlled trial of a computer strategy to increase general practitioner preventive care. *Preventive Medicine* 1999;**29**:478-86.

Borgiel 1999 *{published data only}*

Borgiel AEM, Williams JI, Davis DA, Dunn EV, Hobbs N, Hutchison B, et al. Evaluating the effectiveness of 2 educational interventions on family practice. *Canadian Medical Association* 1999;**8**:965-70.

Brady 1988 *{published data only}*

Brady WJ, Hissa DC, McConnell M, Wones RG. Should physicians perform their own quality assurance audits?. *Journal of General Internal Medicine* 1988;**3**:560-5.

Bregnhoj 2009 *{published data only}*

Bregnhøj L, Thirstrup S, Kristensen MB, Bjerrum L, Sonne J. Combined intervention programme reduces inappropriate

prescribing in elderly patients exposed to polypharmacy in primary care. *European Journal of Clinical Pharmacology* 2009;**65**(2):199-207.

Brown 1994 *{published data only}*

Brown LF, Keily PA, Spencer AJ. Evaluation of a continuing education intervention "Periodontics in General Practice". *Community Dentistry and Oral Epidemiology* 1994;**22**:441-7.

Buffington 1991 *{published data only}*

Buffington J, Bell KM, LaForce FM. A target-based model for increasing influenza immunizations in private practice. *Journal of General Internal Medicine* 1991;**6**:204-9.

Buntinx 1993 *{published data only}*

Buntinx F, Knottnerus JA, Crebolder HF, Seegers T, Essed GG, Schouten H. Does feedback improve the quality of cervical smears? A randomized controlled trial. *British Journal of General Practice* 1993;**43**:194-8.

Canovas 2009 *{published data only}*

Cánovas JJ, Hernández PJ, Botella JJ. Effectiveness of internal quality assurance programmes in improving clinical practice and reducing costs. *Journal of Evaluation in Clinical Practice* 2009;**15**(5):813-9.

Charrier 2008 *{published data only}*

Charrier L, Allochis MC, Cavallo MR, Gregori D, Cavallo F, Zotti CM. Integrated audit as a means to implement unit protocols: a randomized and controlled study. *Journal of Evaluation in Clinical Practice* 2008;**14**(5):847-53.

Chassin 1986 *{published data only}*

Chassin MR, McCue SM. A randomized trial of medical quality assurance. Improving physicians' use of pelvimetry. *JAMA* 1986;**256**:1012-6.

Cheater 2006 *{published data only}*

Cheater FM, Baker R, Reddish S, Spiers N, Wailoo A, Gillies C, et al. Cluster randomized controlled trial of the effectiveness of audit and feedback and educational outreach on improving nursing practice and patient outcomes. *Medical care* 2006;**44**(6):542-51.

Claes 2005 *{published data only}*

Claes N, Buntinx F, Vijgen J, Arnout J, Vermylen J, Fieuws S, et al. The Belgian Improvement Study on Oral Anticoagulation Therapy: a randomized clinical trial. *European Heart Journal* 2005;**26**:2159-65.

Cline 2007 *{published data only}*

Cline D, Ayala C, Caskie D, Ferrario C. Patient specific feedback increases referral of hypertensive emergency department patients: a randomized controlled trial. *Academic Emergency Medicine* 2007;**14**(5):S117.

Cohen 1982 *{published data only}*

Cohen DI, Jones P, Littenberg B, Neuhauser D. Does cost information availability reduce physician test usage? A randomized clinical trial with unexpected findings. *Medical Care* 1982;**20**:286-92.

Curran 2008 *{published data only}*

Curran E, Harper P, Loveday H, Gilmour H, Jones S, Benneyan J, et al. Results of a multicentre randomised controlled trial of statistical process control charts and structured diagnostic tools to reduce ward-acquired meticillin-resistant Staphylococcus aureus: the CHART Project. *Journal of Hospital Infection* 2008;**70**(2):127-35.

Curtis 2005 *{published data only}*

Curtis JR, Olivieri J, Allison JJ, Gaffo A, Juarez L, Kovac SH, et al. A group randomized trial to improve safe use of nonsteroidal anti-inflammatory drugs. *American Journal of Managed Care* 2005;**11**(9):537-43.

Curtis 2007 *{published data only}*

Curtis JR, Westfall AO, Allison J, Becker A, Melton ME, Freeman A, et al. Challenges in improving the quality of osteoporosis care for long-term glucocorticoid users: a prospective randomized trial. *Archives of Internal Medicine* 2007;**167**(6):591-6.

De Almeida Neto 2000 *{published data only}*

Neto ACDA, Benrimoj SI, Kavanagh DJ, Boakes RA. A pharmacy based protocol and training program for non-prescription analgesics. *Journal of Social and Administrative Pharmacy* 2000;**17**(3):183-92.

Eccles 2001 *{published data only}*

Eccles M, Steen N, Grimshaw J, Thomas L, McNamee P, Soutter J, et al. Effect of audit and feedback, and reminder messages on primary-care radiology referrals: a randomised trial. *Lancet* 2001;**357**(9266):1406-9.

Eltayeb 2005 *{published data only}*

Eltayeb IB, Awad AI, Mohamed-Salih MS, Daffa-Alla MA, Ahmed MB, Ogail MA, et al. Changing the prescribing patterns of sexually transmitted infections in the White Nile Region of Sudan. *Sexually Transmitted Infections* 2005;**81**:426-7.

Everett 1983 *{published data only}*

Everett GD, deBlois CS, Chang PF, Holets T. Effect of cost education, cost audits, and faculty chart review on the use of laboratory services. *Archives of Internal Medicine* 1983;**143**:942-4.

Fairbrother 1999 *{published data only}*

Fairbrother G, Hanson KL, Friedman S, Butts GC. The impact of physician bonuses, enhanced fees, and feedback on childhood immunization coverage rates. *American Journal of Public Health* 1999;**89**(2):171-5.

Ferguson 2003 *{published data only}*

Ferguson TB, Peterson ED, Coombs LP, Eiken MC, Carey ML, Grover FL, et al. Use of contiouous quality improvement to increase use of process measures in patients undergoing coronary artery bypass graft surgery. *JAMA* 2003;**290**(1):49-56.

Filardo 2009 *{published data only}*

Filardo G, Nicewander D, Herrin J, Edwards J, Galimbertti P, Tietze M, et al. A hospital-randomized controlled trial of a formal quality improvement educational program in rural and

small community Texas hospitals: one year results. *International Journal for Quality in Health Care* 2009;**21**:225-32.

Foster 2007 *{published data only}*

Foster JM, Hoskins G, Smith B, Lee AJ, Price D, Pinnock H. Practice development plans to improve the primary care management of acute asthma: randomised controlled trial. *BMC Family Practice* 2007;**8**:23.

Foy 2004 *{published data only}*

Foy R, Penney GC, Grimshaw JM, Ramsay CR, Walker AE, MacLennan G, et al. A randomised controlled trial of a tailored multifaceted strategy to promote implementation of a clinical guideline on induced abortion care. *BJOG: an International Journal of Obstetrics & Gynaecology* 2004;**111**(7):726-33.

Frijiling 2002 *{published data only}*

Frijling BD, Lobo CM, Hulscher MEJL, Akkarmans RP, Braspenning JCC, Prins A, et al. Multifaceted support to improve clinical decision making in diabetes care: a randomized controlled trial in general practice. *Diabetic Medicine* 2002;**19**:836-42.

Frijiling 2003 *{published data only}*

Frijling BD, Lobo CM, Hulscher MEJL, Akkarmans RP, Van Drenth BB, Prins A, et al. Intensive support to improve clinical decision making in cardiovascular care: a randomised controlled trial in general practice. *Quality and Safety in Health Car* 2003;**12**:181-7.

Gama 1992 *{published data only}*

Gama R, Nightingale PG, Broughton PM, Peters M, Ratcliffe JG, Bradby GV, et al. Modifying the request behaviour of clinicians. *Journal of Clincial Pathology* 1992;**45**(8):742-3.

Gehlbach 1984 *{published data only}*

Gehlbach SH, Wilkinson WE, Hammond WE, Clapp NE, Finn AL, Taylor WJ, et al. Improving drug prescribing in a primary care practice. *Medical Care* 1984;**22**:193-201.

Goff 2003 *{published data only}*

Goff DC, Gu L, Cantley LK, Sheedy DJ, Cohen SJ. Quality of care for secondary prevention for patients with coronary heart disease: Results of the hastening the effective application of research through technology (HEART) trial. *American Heart Journal* 2003;**146**(6):1045-151.

Grady 1997 *{published data only}*

Grady KE, Lemkau JP, Lee NR, Caddell C. Enhancing mammography referral in primary care. *Preventive Medicine* 1997;**26**:791-800.

Guadagnoli 2000 *{published data only}*

Guadagnoli E, Soumerai SB, Gurwitz JH, Borbas C, Shapiro CL, Weeks JC, et al. Improving discussion of surgical treatment options for patients with breast cancer: local medical opinion leaders versus audit and performance feedback. *Breast Cancer Research and Treatment* 2000;**61**(2):171-5.

Gullion 1988 *{published data only}*

Gullion DS, Tschann JM, Adamson TE, Coates TJ. Management of hypertension in private practice: a randomized controlled trial in continuing medical education. *The Journal of Continuing Education in the Health Professions* 1988;**8**:239-55.

Hayes 2001 *{published data only}*

Hayes R, Bratzler D, Armour B, Moore l. Comparison of an enhanced versus written feedback model on the management of Medicare inpatients with venous thrombosis. *Joint Commission Journal on Quality Improvement* 2001;**27**(3):155-68.

Hayes 2002 *{published data only}*

Hayes RP, Baker DW, Luthi JC, Baggett RL, McClellan W, FitzGerald D, et al. The effect of external feedback on the management of medicare inpatients with congestive heart failure. *American Journal of Medical Quality* 2002;**17**:225-35.

Heller 2001 *{published data only}*

Heller RF, DEste C, Lim LL, OConnel RL, Powell H. Randomised controlled trial to change hospital management of unstable angina. *Medical Journal of Australia* 2001;**175**(5):217-21.

Hemminiki 1992 *{published data only}*

Hemminiki E, Teperi J, Tuominen K. Need for and influence or feedback from the Finnish birth register to data providers. *Quality Assurance in Health Care* 1992;**4**(2):133-9.

Hendryx 1998 *{published data only}*

Hendryx MS, Fieselmann JF, Bock MJ, Wakefield DS, Helms CM, Bentler SE. Outreach education to improve quality of rural icu care. *American Journal of Respiratory Critical Care Medicine* 1998;**158**:418-23.

Herbert 2004 *{published data only}*

Herbert CP, Wright JM, Maclure M, Wakefield J, Dormuth C, Brett-MacLean P, et al. Better Prescribing Project: a randomized controlled trial of the impact of case-based educational modules and personal prescribing feedback on prescribing for hypertension in primary care. *Family Practice* 2004;**21**:575-81.

Herrin 2006 *{published data only}*

Herrin J, Nicewander DA, Hollander PA, Couch CE, Winter FD, Haydar ZR, et al. Effectiveness of diabetes resource nurse case management and physician profiling in a fee-for-service setting: a cluster randomized trial. *Proceedings (Baylor University Medical Centre)* 2006;**19**(2):95-102.

Hershey 1986 *{published data only}*

Hershey CO, Porter DK, Breslau D, Cohen DI. Influence of simple computerized feedback on prescription charges in an ambulatory clinic. A randomized clinical trial. *Medical Care* 1986;**24**:472-81.

Hershey 1988 *{published data only}*

Hershey CO, Goldberg HI, Cohen DI. The effect of computerized feedback coupled with a newsletter upon outpatient prescribing charges. A randomized controlled trial. *Medical Care* 1988;**26**(1):88-94. [MEDLINE: 88093018]

Hillman 1998 *{published data only}*

Hillman AL, Ripley K, Goldfarb N, Nuamah I, Weiner J, Lusk E. Physician financial incentives and feedback: Failure to increase cancer screening in medicaid managed care. *American Journal of Public Health* 1998;**88**(11):1698-701.

Hillman 1999 *{published data only}*

Hillman AL, Ripley K, Goldfarb N, Weiner J, Nuamah I, Lusk E. The use of physician financial incentives and feedback to improve pediatric preventive care in Medicaid care. *Pediatrics* 1999;**104**(4):931-5.

Holm 1990 *{published data only}*

Holm M. Intervention against long-term use if hypnotics/ sedatives in general practice. *Scandinavian Journal of Primary Health Care* 1990;**8**:113-7.

Hux 1999 *{published data only}*

Hux JE, Melady MP, DeBoer D. Confidential prescriber feedback and education to improve antibiotic use in primary care: a controlled trial. *Canadian Medical Association* 1999;**161**:388-92.

Kahan 2009 *{published data only}*

Kahan NR, Kahan E, Waitman DA, Kitai E, Chintz DP. The tools of an evidence-based culture: implementing clinical-practice guidelines in an Israeli HMO. *Academic Medicine* 2009;**84**(9):1217-25.

Kerry 2000 *{published data only}*

Kerry S, Oakeshott P, Dundas D, Williams J. Influence of postal distribution of the royal college of radiologists guidelines, together with feedback on radiological referral rates, on x-ray referrals from general practice: a randomized controlled trial. *Family Practice* 2000;**17**(1):46-52.

Kiefe 2001 *{published data only}*

Kiefe CI, Allison JJ, Williams OD, Person SD, Weaver MT, Weissman NW. Improving quality improvement using achievable benchmarks for physician feedback: a randomized controlled trial. *JAMA* 2001;**285**(22):2871-9.

Kim 1999 *{published data only}*

Kim CS, Kristopaitis RJ, Stone E, Pelter M, Sandhu M, Weingarten SR. Physician education and report cards: do they make the grade? Results from a randomized controlled trial. *The American Journal of Medicine* 1999;**107**:556-60.

Kinsinger 1998 *{published data only}*

Kinsinger LS, Harris R, Qaqish B, Strecher V, Kaluzny A. Using an office system intervention to increase breast cancer screening. *Journal of General Internal Medicine* 1998;**13**:507-14.

Kogan 2003 *{published data only}*

Kogan JR, Reynolds EE, Shea JA. Effectiveness of report cards based on chart audits of residents adherence to practice guidelines on practice performance: a randomized controlled trial. *Teaching and Learning in Medicine* 2003;**15**(1):25-30.

Kritchevsky 2008 *{published data only}*

Kritchevsky SB, Braun BI, Bush AJ, Bozikis MR, Kusek L, Burke JP, et al. TRAPE Study Group. The effect of a quality improvement collaborative to improve antimicrobial prophylaxis in surgical patients: a randomized trial. *Annals of Internal Medicine* 2008;**149**(7):472-80, W89-93.

Lagerløv 2000 *{published data only}*

Lagerløv P, Loeb M, Andrew M, Hjortdahl P. Improving doctors' prescribing behaviour through reflection on guidelines and prescription feedback: a randomised controlled study. *Quality in Health Care* 2000;**9**:159-65.

Laskshminarayan 2010 *{published data only}*

Lakshminarayan K, Borbas C, McLaughlin B, Morris NE, Vazquez G, Luepker RV, et al. A cluster-randomized trial to improve stroke care in hospitals. *Neurology* 2010;**74**:1634-42.

Linn 1980 *{published data only}*

Linn BS. Continuing medical education. Impact on emergency room burn care. *JAMA* 1980;**244**:565-70.

Lobach 1996 *{published data only}*

Lobach DF. Electronically distributed computer-generated feedback enhances the use of a computarized practice guidelines. Proceedings/AMIA Annual Fall symposium. 1996:493-7.

Lomas 1991 *{published data only}*

Lomas J, Enkin M, Anderson GM, Hannah WJ, Vayda E, Singer J. Opinion leaders vs audit and feedback to implement practice guidelines. Delivery after previous cesarean section. *JAMA* 1991;**265**:2202-7.

Mainous 2000 *{published data only}*

Mainous AG, Hueston WJ, Love MM, Evans ME, Finger R. An evaluation of statewide strategies to reduce antibiotic overuse. *Family Medicine* 2000;**32**(1):22-9.

Martin 1980 *{published data only}*

Martin AR, Wolf MA, Thibodeau LA, Dzau V, Braunwald E. A trial of two strategies to modify the test-ordering behavior of medical residents. *New England Journal of Medicine* 1980;**303**:1330-6.

Marton 1985 *{published data only}*

Marton KI, Tul V, Sox HC Jr. Modifying test-ordering behavior in the outpatient medical clinic. A controlled trial of two educational interventions. *Archives of Internal Medicine* 1985;**145**:816-21.

Mayer 1998 *{published data only}*

Mayer JA, Eckhardt L, Stepanski BM, Sallis JF, Elder JP, Slymen DJ, et al. Promoting skin cancer prevention counseling. *American Journal for Public Health* 1998;**88**(7):1096-9.

McAlister 1986 *{published data only}*

McAlister NH, Covvey HD, Tong C, Lee A, Wigle ED. Randomised controlled trial of computer assisted management of hypertension in primary care. *BMJ* 1986;**293**:670-4.

McCartney 1997 *{published data only}*

McCartney P, Macdowall W, Thorogood M. A randomised controlled trial of feedback to general, practitioners of their prophylactic aspirin prescribing. *BMJ* 1997;**315**:35-6.

McClellan 2003 *{published data only}*

McClellan WM, Millman L, Presley R, Couzins J, Flanders WD. Improved diabetes care by primary care physicians: results of a group-randomized evaluation of the Medicare Health Care Quality Improvement Program (HCQIP). *Journal of Clinical Epidemiology* 2003;**56**:1210-7.

McClellan 2004 *{published data only}*

McClellan WM, Hodgin E, Pastan S, McAdams L, Soucie M. A randomized evaluation of two health care quality improvement program (HCQIP) interventions to improve the adequacy of hemodialysis care of ESRD patients: feedback alone versus intensive intervention. *Journal of the American Society of Nephrology* 2004;**15**:754-60.

McConnell 1982 *{published data only}*

McCollell TS, Cushing AH, Healy JL, McIlvenna PA, Skipper BJ. Physician behavior modification using claims data: tetracycline for upper respiratory infection. *The Western Journal of Medicine* 1982;**137**(5):448-50.

Millard 2008 *{published data only}*

Millard FB, Thistlethwaite J, Spagnolo C, Kennedy RL, Baune BT. Dementia diagnosis: A pilot randomised controlled trial of education and IT audit to assess change in GP dementia documentation. *Australian Journal of Primary Health* 2008;**14**:141-9.

Mitchell 2005 *{published data only}*

Mitchell E, Sullivan F, Grimshaw JM, Donnan PT, Watt G. Improving management of hypertension in general practice: a randomised controlled trial of feedback derived from electronic patient data. *British Journal of General Practice* 2005;**55**(511):94-101.

Moher 2001 *{published data only}*

Moher M, Yudkin P, Wright L, Turner R. Cluster randomised controlled trial to compare three methods of promoting secondary prevention of coronary hearth disease in primary care. *BMJ* 2001;**322**(7298):1338.

Mold 2008 *{published data only}*

Mold JW, Aspy CA, Nagykaldi Z, Oklahoma Physicians Resource/ Research Network. Implementation of evidence-based preventive services delivery processes in primary care: an Oklahoma Physicians Resource/Research Network (OKPRN) study. *Journal of the American board of family medicine* 2008;**21**(4):334-44.

Naughton 2007 *{published data only}*

Naughton C, Feely J, Bennett K. A clustered randomized trial of the effects of feedback using academic detailing compared to postal bulletin on prescribing of preventative cardiovascular therapy. *Family Practice* 2007;**24**(5):475-80.

Naughton 2009 *{published data only}*

Naughton C, Feely J, Bennett K. A RCT evaluating the effectiveness and cost-effectiveness of academic detailing versus postal prescribing feedback in changing GP antibiotic prescribing. *Journal of Evaluation in Clinical Practice* 2009;**15**(5):807-12.

Nilsson 2001 *{published data only}*

Nilsson G, Hjemdal P, Hassler A, Vitols S, Wallen NH, Krakau I. Feedback on prescribing rate combined with problem-oriented pharmacotherapy education as a model to improve prescribing among general practitioners. *European Journal of Clinical Pharmacology* 2001;**56**(11):843-8.

Norton 1985 *{published data only}*

Norton PG, Dempsey LJ. Self-audit: its effect on quality of care. *Journal of Family Practice* 1985;**21**:289-91.

O'Connell 1999 *{published data only}*

O´Connell DL, Henry D, Tomlins R. Randomised controlled trial of effect of feedback on general practitioners prescribing in Australia. *BMJ* 1999;**318**:507-11.

O'Connor 2009 *{published data only}*

O'Connor PJ, Sperl-Hillen J, Johnson PE, Rush WA, Crain AL. Customized feedback to patients and providers failed to improve safety or quality of diabetes care: a randomized trial. *Diabetes Care* 2009;**32**(7):1158-63.

Ornstein 2004 *{published data only}*

Ornstein S, Jenkins RG, Nietert PJ, Feifer C, Roylance LF, Nemeth L, et al. A multimethod quality improvement intervention to improve preventive cardiovascular care: a cluster randomized trial. *Annals of Internal Practice* 2004;**141**(7):523-32.

Palmer 1985 *{published data only}*

Palmer RH, Louis TA, Hsu LN, Peterson HF, Rothrock JK, Strain R, et al. A randomized controlled trial of quality assurance in sixteen ambulatory care practices. *Medical Care* 1985;**23**:751-70.

Phillips 2005 *{published data only}*

Phillips LS, Ziemer DC, Doyle JP, Barnes CS, Kolm P, Branch WT, et al. An endocrinologist-supported intervention aimed at providers improves diabetes management in a primary care site: improving primary care of African Americans with diabetes (IPCAAD) 7. *Diabetes Care* 2005;**28**:2352-60.

Pimlott 2003 *{published data only}*

Pimlott NJG, Hux JE, Wilson LM, Kahan M, Li C, Rosser WW. Educating physicians to reduce benzodiazepine use by elderly patients: a randomized controlled trial. *Canadian Medical Association Journal* 2003;**168**:835-9.

Quinley 2004 *{published data only}*

Quinley JC, Shih A. Improving physician coverage of pneumcoccal vaccine: a randomized trial of telephone intervention. *Journal of Community Health* 2004;**29**:103-15.

Raasch 2000 *{published data only}*

Raasch BA, Hays R, Buettner PG. An educational intervention to improve diagnosis and management of suspicious skin lesions. *The Journal of Continuing Education in the Health Professions* 2000;**20**:39-51.

Rantz 2001 *{published data only}*

Rantz MJ, Popejoy L, Petroski GF, Madsen RW, Mehr DR, Zwygart-Stauffacher M, et al. Randomized clinical trial of quality improvement intervention in nursing homes. *The Gerontologist* 2001;**41**(4):525-38.

Rask 2001 *{published data only}*

Rask K, Kohler SA, Wells KJ, Williams JA, Diamond CC. Performance improvement interventions to improve delivery of screening services in diabetes care. *Journal of Clinical Outcomes Management* 2001;**8**:23-9.

Robling 2002 *{published data only}*

Robling MR, Houston HL, Kinnersley P, Hourihan MD, Cohen DR, Hale J, et al. General practitioners' use of magnetic resonance imaging: an open randomized trial comparing telephone and written requests and an open randomized controlled trial of different methods of local guideline dissemination. *Clinical Radiology* 2002;**57**(5):402-7.

Ruangkanchanasetr 1993 *{published data only}*

Ruangkanchanastr S. Laboratory investigation utilization in pediatric out-patient department ramathibodi hospital. *Journal of the Medical Association of Thailand* 1993;**76**:194-9.

Rubin 2001 *{published data only}*

Rubin GL, Schofield WN, Dean MG, Shakeshaft AP. Appropriateness of red blood cell transfusions in major urban hospitals and effectiveness of an intervention. *The Medical Journal of Australia* 2001;**175**:354-8.

Rust 1999 *{published data only}*

Rust CT, Sisk FA, Kuo AR, Smith J, Miller R, Sullivan KM. Impact of resident feedback on immunization outcomes. *Archives of Pediatrics & Adolescent Medicine* 1999;**153**:1165-9.

Sandbaek 1999 *{published data only}*

Sandbaek A, Kragstrup J. Randomized controlled trial of the effect of medical audit on aids prevention in general practice. *Family Practice* 1999;**16**:510-4.

Sauaia 2000 *{published data only}*

Sauaia A, Ralston D, Schluter WW, Marciniak TA, Havranek EP, Dunn TR. Influencing care in acute myocardial infarction: a randomized trial comparing 2 types of intervention. *American Journal of Medical Quality* 2000;**15**:197-206.

Schectman 1995 *{published data only}*

Schectman JM, Kanwal NK, Schroth WS, Elinsky EG. The effect of an education and feedback intervention on group-model and network-model health maintenance organization physician prescribing behavior. *Medical Care* 1995;**33**:139-44.

Schectman 2003 *{published data only}*

Schectman JM, Schroth WS, Verme D, Voss JD. Randomized controlled trial of education and feedback for implementation of guidelines for acute low back pain. *Journal of General Internal Medicine* 2003;**18**:773-80.

Schneider 2008 *{published data only}*

Schneider A, Wensing M, Biessecker K, Quinzler R, Kaufmann-Kolle P, Szecsenyi J. Impact of quality circles for improvement of asthma care: results of a randomized controlled trial. *Journal of Evaluation in Clinical Practice* 2008;**14**(2):185-90.

Scholes 2006 *{published data only}*

Scholes D, Grothaus L, McClure J, Reid R, Fishman P, Sisk C, et al. A randomized trial of strategies to increase chlamydia screening in young women. *Preventive Medicine* 2006;**43**(4):343-50.

Sinclair 1982 *{published data only}*

Sinclair C, Frankel M. The effect of quality assurance activities on the quality of mental health services. *Quality Review Bulletin* 1982;**8**(7):7-15.

Siriwardena 2002 *{published data only}*

Siriwardena AN, Rashid A, Johnson MRD, Dewey ME. Cluster randomised controlled trial of an educational outreach visit to improve influenza and pneumococcal immunisation rates in primary care. *The British Journal of General Practice* 2002;**52**:735-40.

Smith 1998 *{published data only}*

Smith DK, Shaw RW, Slack J, Marteau TM. Training obstetricians and midwives to present screening tests evaluation of two brief interventions. *Prenatal Diagnosis* 1995;**15**:317-24.

Socolar 1998 *{published data only}*

Socolar RR, Raines B, Chen-Mok M, Runyan DK, Green C, Paterno S. Intervention to improve physician documentation and knowledge of child sexual abuse: a randomized, controlled trial. *Pediatrics* 1998;**101**(5):817-24.

Solomon 2004 *{published data only}*

Solomon DH, Katz JN, La Tourette AM, Coblyn JS. Multifaceted intervention to improve rheumatologists' management of glucocorticoid-induced osteoporosis: a randomized controlled trial. *Arthritis and Rheumatism* 2004;**51**(3):383-7.

Sommers 1984 *{published data only}*

Sommers LS, Sholtz R, Shepherd RM, Starkweather DB. Physician involvement in quality assurance. *Medical Care* 1984;**22**:1115-38.

Soumerai 1998 *{published data only}*

Soumerai SB, McLaughlin TJ, Gurwitz JH, Guadagnoli E, Hauptman PJ, Borbas C, et al. Effect of local medical opinion leaders on quality of care for acute myocardial infarction: a randomized controlled trial. *JAMA* 1998;**279**(17):1358-63.

Svetkey 2009 *{published data only}*

Svetkey LP, Pollak KI, Yancy WS Jr, Dolor RJ, Batch BC, Samsa G, et al. Hypertension improvement project: randomized trial of

quality improvement for physicians and lifestyle modification for patients. *Hypertension* 2009;**54**(6):1226-33.

Søndergaard 2002 *{published data only}*

Søndergaard J, Adersen M, Vach K, Kragstrup J, Maclure M, Gram LF. Detailed postal feedback about prescribing to asthma patients combined with a guideline statement showed no impact: a randomised controlled trial. *European Journal of Clinical Pharmacology* 2002;**58**:127-32.

Søndergaard 2003 *{published data only}*

Søndergaard J, Andersen M, Støvring H, Kragstrup J. Mailed prescribed feedback in addition to a clinical guideline has no impact: a randomised, controlled trial. *Scandinavian Journal of Primary Health Care* 2003;**21**:47-51.

Søndergaard 2006 *{published data only}*

Søndergaard J, Hansen DG, Aarslev P, Munck AP. A multifaceted intervention according to the Audit Project Odense method improved secondary prevention of ischemic heart disease: a randomised controlled trial. *Family Practice* 2006;**23**(2):198-202.

Thomas 2006 *{published data only}*

Thomas RE, Croal BL, Ramsay C, Eccles M, Grimshaw J. Effect of enhanced feedback and brief educational reminder messages on laboratory test requesting in primary care: a cluster randomised trial. *Lancet* 2006;**367**:1990-6.

Thomas 2007 *{published data only}*

Thomas KG, Thomas MR, Stroebel RJ, McDonald FS, Hanson GJ, Naessens JM, et al. Use of a registry-generated audit, feedback, and patient reminder intervention in an internal medicine resident clinic--a randomized trial. *Journal of General Internal Medicine* 2007;**22**:1740-4.

Tierney 1986 *{published data only}*

Tierney WM, Hui SL, McDonald CJ. Delayed feedback of physician performance versus immediate reminders to perform preventive care. Effects on physician compliance. *Medical Care* 1986;**24**(8):659-66.

Tu 2009 *{published data only}*

Tu JV, Donovan LR, Lee DS, Wang JT, Austin PC, Alter DA, et al. Effectiveness of public report cards for improving the quality of cardiac care: the EFFECT study: a randomized trial. *JAMA* 2009;**302**(21):2330-7.

Van den Hombergh 1999 *{published data only}*

Van den Hombergh P, Grol R, Van den Hoogen HJM, Van den Bosch WJHM. Practice visits as a tool in quality improvement: mutual visits and feedback by peers compared with visits and feedback by non-physician observers. *Quality in Health Care* 1999;**8**:161-6.

Van der Weijden 1999 *{published data only}*

Van der Weijden T, Grol RP, Knottinerus JA. Feasibility of a national cholestrol guideline in daily practice. A randomized controlled trial in 20 general practices. *International Journal for Quality in Health Care* 1999;**11**(2):131-7.

Veninga 1999 *{published data only}*

Lundborg CS, Wahlström, Oke T, Tomson G, Diwan V. Influencing prescribing for urinary tract infection and asthma in primary care in Sweden: a randomized controlled trial of an interactive educational intervention. *Journal of Clinical Epidemiology* 1999;**52**(8):801-12.

Veninga CCM, Denig P, Zwaagstra R, Haaijer-Ruskamp FM. Improving drug treatment in general practice. *Journal of Clinical Epidemiology* 2000;**53**:762-72.

* Veninga CCM, Lagerløv P, Wahlstöm R, Muskova M, Denig P, Berkhof J, et al. Evaluating an educational intervention to improve the treatment of asthma in four European countries. *American Journal of Respiratory and Critical Care Medicine* 1999;**160**:1254-62.

Veninga N. Improving prescribing in general practice. Thesis, Rijksuiversiteit Groningen 2000.

Verstappen 2003 *{published data only}*

Verstappen WHJM, Van der Weijden T, Sijbrandij J, Smeele J, Hermsen J, Grimshaw J, et al. Effect of a practice-based strategy on test ordering performance of primary care physicians. *JAMA* 2003;**289**:2407-12.

Verstappen 2004 *{published data only}*

Verstappen WHJM, Van der Weijden T, Dubois WI, Smeele I, Hermsen J, Tan FES, et al. Improving test ordering in primary care: the added value of a small-group quality improvement strategy compared with classic feedback only. *Annals of Family Medicine* 2004;**2**:569-75.

Vingerhoets 2001 *{published data only}*

Vingerhoets B, Wensing M, Grol R. Feedback of patients' evaluations of general practice care: a randomised trial. *Quality in Health Care* 2001;**10**:224-8.

Wadland 2007 *{published data only}*

Wadland WC, Holtrop JS, Weismantel D, Pathak PK, Fadel H, Powell J. Practice-based referrals to a tobacco cessation quit line: assessing the impact of comparative feedback vs general reminders. *Annals of Family Medicine* 2007;**5**(2):135-42.

Wahlström 2003 *{published data only}*

Wahlström R, Kounnavong S, Sisounthone B, Phanyanouvong A, Southammavong T, Eriksson B, et al. Effectiveness of feedback for improving case management of malaria, diarrhoea and pneumonia - a randomized controlled trial at provincial hospitals in Lao PDR. *Tropical Medicine and International Health* 2003;**8**:901-9.

Ward 1996 *{published data only}*

Ward A, Kamien M, Mansfield F, Fatovich B. Educational feedback in management of diabetes in general practice. *Education for General Practice* 1996;**7**:142-50.

Weitzman 2009 *{published data only}*

Weitzman S, Greenfield S, Billimek J, Hava T, Schvartzman P, Yehiel E, et al. Improving combined diabetes outcomes by adding a simple patient intervention to physician feedback:

a cluster randomized trial. *Israeli Medical Association Journal* 2009;**11**:719-24.

Winickoff 1984 *{published data only}*

Winickoff RN, Coltin KL, Morgan MM, Buxbaum RC, Barnett GO. Improving physician performance through peer comparison feedback. *Medical Care* 1984;**22**:527-34.

Winkens 1995 *{published data only}*

Winkens RA, Pop P, Bugter-Maessen AM, Grol RP, Kester AD, Beusmands GH, et al. Randomised controlled trial of routine individual feedback to improve rationality and reduce numbers of test requests. *Lancet* 1995;**345**:498-502.

Wones 1987 *{published data only}*

Wones RG. Failure of low-cost audits with feedback to reduce laboratory test utilization. *Medical Care* 1987;**25**:78-82.

Ziemer 2006 *{published data only}*

Ziemer DC, Doyle JP, Barnes CS, Branch WTJ, Cook CB, El Kebbi IM, et al. An intervention to overcome clinical inertia and improve diabetes mellitus control in a primary care setting: Improving Primary Care of African Americans with Diabetes (IPCAAD) 8. *Archives of Internal Medicine* 2006;**166**:507-13.

Zwar 1999 *{published data only}*

Zwar N, Wolk J, Gordon J, Fisher RS, Kehoe L. Influencing antibiotic prescribing in general practice: a trial of prescriber feedback and management guidelines. *Family Practice* 1999;**16**(5):495-500.

References to studies excluded from this review

Aittasalo 2006 *{published data only}*

Aittasalo M, Miilunpalo S, Kukkonen-Harjula K, Pasanen M. A randomized intervention of physical activity promotion and patient self-monitoring in primary health care. *Preventive Medicine* 2006;**42**:40-6.

Allard 2001 *{published data only}*

Allard J, Hébert R, Rioux M, Asselin J, Voyer L. Efficacy of a clinical medication review on the number of potentially inappropriate prescriptions prescribed for community-dwelling elderly people. *Canadian Medical Association Journal* 2001;**164**:1291-6.

Allison 2005 *{published data only}*

Allison JJ, Kiefe CI, Wall T, Casebeer L, Ray MN, Spettell CM, et al. Multicomponent Internet continuing medical education to promote chlamydia screening.. *American Journal of Preventive Medicine* 2005;**28**:285-90.

Althabe 2008 *{published data only}*

Althabe F, Buekens P, Bergel E, Belizan JM, Campbell MK, Moss N, et al. A behavioral intervention to improve obstetrical care. *New England Journal of Medicine* 2008;**358**(18):1929-40.

Anderson 1996 *{published data only}*

Anderson JF, McEwan KL, Hrudey WP. Effectiveness of notification and group education in modifying prescribing of regulated analgesics. *Canadian Medical Association Journal* 1996;**154**:31-9.

Anonymous I 1990 *{published data only}*

Anonymous. North of England study of standards and performence in general practice. University of Newcastle, report nr 40 and report nr 50.

Aspy 2008 *{published data only}*

Aspy CB, Enright M, Halstead L, Mold JW, Oklahoma Physicians Resource/Research Network. Improving mammography screening using best practices and practice enhancement assistants: an Oklahoma Physicians Resource/Research Network (OKPRN) study. *Journal of the American Board of Family Medicine: JABFM* 2008;**21**(4):326-33.

Ballard 2002 *{published data only}*

Ballard DJ, Nicewander D, Skinner C. Health care provider quality improvement organization medicare data sharing: A diabetes quality improvement initiative. Proceedings of AMIA Symposium. 2002:22-5.

Belcher 1990 *{published data only}*

Belcher DV. Implementing preventive services success and failure in an outpatient trial. *Archives of Internal Medicine* 1990;**150**:2533-41.

Bertoni 2009 *{published data only}*

Bertoni AG, Bonds DE, Chen H, Hogan P, Crago L, Rosenberger E, et al. Impact of a multifaceted intervention on cholesterol management in primary care practices: guideline adherence for heart health randomized trial. *Archives of Internal Medicine* 2009;**169**(7):678-86.

Berwick 1986 *{published data only}*

Berwick DM, Coltin KL. Feedback reduces test use in a health maintenance organization. *JAMA* 1986;**255**(11):1450-4. [MEDLINE: 86144192]

Billi 1987 *{published data only}*

Billi JE, Hejna GF, Wolf FM, Shapiri LR, Stross JK. The effects of a cost-education program on hospital charges. *Journal of General Internal Medicine* 1987;**2**:306-11.

Bindels 2003 *{published data only}*

Bindels R, Hasman A, Kester AD, Talmon JL, De Clercq PA, Winkens RA. The efficacy of an automated feedback system for general practitioners. *Informatics in Primary Care* 2003;**11**:69-74.

Bischoff 2000 *{published data only}*

Bischoff WE, Reynolds TM, Sessler CN, Edmond MB, Wenzel RP. Handwashing compliance by health care workers: The impact of introducing an accessible, alcohol-based hand antiseptic. *Archives of Internal Medicine* 2000;**160**:1017-21.

Bonds 2009 *{published data only}*

Bonds DE, Hogan PE, Bertoni AG, Chen H, Clinch CR, Hiott AE, et al. A multifaceted intervention to improve blood pressure control: The Guideline Adherence for Heart Health (GLAD) study. *American Heart Journal* 2009;**157**(2):278-84.

Bonetti 2005 *{published data only}*

Bonetti D, Eccles M, Johnston M, Steen N, Grimshaw J, Baker R, et al. Guiding the design and selection of interventions to influence the implementation of evidence-based practice: an experimental simulation of a complex intervention trial. *Social Science & Medicine* 2005;**60**:2135-47.

Brand 2005 *{published data only}*

Brand S, Härter M, Sitta P, Van Calker D, Menke R, Heindl A, et al. Data supporting quality circle management of inpatient depression treatment. *Nervenarzt* 2005;**76**:865-6.

Britton 1991 *{published data only}*

Britton ML, Lurvey PL. Impact of medication profile review on prescribing in a general medicine clinic. *American Journal of Hospital Pharmacy* 1991;**48**:265-70.

Brown 1988 *{published data only}*

Brown RL. Evaluation of a continuing medical education program for primary care physcicians on the management of alcoholism. *Journal of Medical Education* 1988;**63**:482-4.

Buekens 1993 *{published data only}*

Buekens P, Boutsen M, Kittel F, Vandenbussche P, Dramaix M. Does awareness of rates of obstetric interventions change practice?. *BMJ* 1993;**306**:623.

Bunting 2004 *{published data only}*

Bunting PS, Van Walraven C. Effect of a controlled feedback intervention on laboratory test ordering by community physicians. *Clinical Chemistry* 2004;**50**:321-6.

Campbell 2006 *{published data only}*

Campbell E, Walsh RA, Sanson-Fisher R, Burrows S, Stojanovski E. A group randomised trial of two methods for disseminating a smoking cessation programme to public antenatal clinics: effects on patient outcomes. *Tobacco Control* 2006;**15**(2):97-102.

Carney 1992 *{published data only}*

Carney PA, Dietrich AJ, Keller A, Landgraf J, O'Connor GT. Tools, teamwork and tenacity: An office system for cancer prevention. *The Journal of Family Practice* 1992;**35**:388-94.

Chin 2007 *{published data only}*

Chin MH, Drum ML, Guillen M, Rimington A, Levie JR, Kirchhoff AC, et al. Improving and sustaining diabetes care in community health centers with the health disparities collaboratives. *Medical Care* 2007;**45**:1135-43.

Chowdhury 2007 *{published data only}*

Chowdhury AKA, Khan OF, Matin MA, Begum K, Galib MA. Effect of standard treatment guidelines with or without prescription audit on prescribing for acute respiratory tract infection (ARI) and diarrhoea in some thana health complexes (THCs) of Bangladesh. *Bangladesh Medical Research Council Bulletin* 2007;**33**:21-30.

Cleveringa 2008 *{published data only}*

Cleveringa FGW, Gorter KJ, Van Donk MD, Rutten GEHM. Combined task delegation, computerized decision support, and feedback improve cardiovascular risk for type 2 diabetic patients. *Diabetes Care* 2008;**31**(12):2273-5.

Cohen 1996 *{published data only}*

Cohen MM, Rose DK, Yee DA. Changing anesthesiologists' practice patterns. Can it be done?. *Anesthesiology* 1996;**85**:260-9.

Colón-Emeric 2007 *{published data only}*

Colón-Emeric CS, Lyles KW, House P, Levine DA, Schenck AP, Allison J, et al. Randomized trial to improve fracture prevention in nursing home residents. *The American Journal of Medicine* 2007;**120**(10):886-92.

Cope 1986 *{published data only}*

Cope DW, Linn LS, Leake BD, Barrett PA. Modification of residents' behavior by preceptor feedback of patient satisfaction. *Journal of General Internal Medicine* 1986;**1**:394-8.

Cranney 1999 *{published data only}*

Cranney M, Barton S, Walley T. Addressing barriers to change: an RCT of practice-based education to improve the management of hypertension in the elderly. *The British Journal of General Practice* 1999;**49**:522-6.

Crits-Christoph 2010 *{published data only}*

Crits-Christoph P, Ring-Kurtz S, McClure B, Temes C, Kulaga A, Gallop R, et al. A randomized controlled study of a web-based performance improvement system for substance abuse treatment providers. *Journal of Substance Abuse Treatment* 2010;**38**:251-62.

Crotty 2004 *{published data only}*

Crotty M, Whitehead C, Rowett D, Halbert J, Weller D, Finucane P, et al. An outreach intervention to implement evidence based practice in residential care: a randomized controlled trial. *BMC Health Services Research* 2004;**4**:6.

Curtis 2009 *{published data only}*

Curtis JR, Nielsen E, Treece P, Downey L, Dotolo D, Shannon S, et al. Critical Care Medicine. 2009; Conference:A388.

Del Mar 1998 *{published data only}*

Del Mar CB, Lowe JB, Adkins P, Arnold E, Baade P. Improving general practitioner clinical records with a quality assurance minimal intervention. *British Journal of General Practice* 1998;**48**:1307-11.

Denton 2001 *{published data only}*

Denton GD, Smith J, Faust J, Holmboe E. Comparing the efficacy of staff versus housestaff instruction in an intervention to improve hypertension management. *Academic Medicine* 2001;**76**:1257-60.

De Silva 1994 *{published data only}*

De Silva M, Abrahanson G. Does medical audit change practice?. *Transfusion Science* 1994;**15**:277.

Dickinson 1981 *{published data only}*

Dickinson JC, Warshaw GA, Gehlbach SH, Bobula JA, Muhlbaier LH, Parkerson GR Jr. Improving hypertension control:

impact of computer feedback and physician education. *Medical Care* 1981;**19**:843-54.

Doherty 2006 *{published data only}*

Doherty SR, Jones PD. Use of an 'evidence-based implementation' strategy to implement evidence-based care of asthma into rural district hospital emergency departments. *Rural and Remote Health* 2006;**6**:529.

Doherty 2007 *{published data only}*

Doherty SR, Jones PD, Davis L, Ryan NJV, Treeve V. Evidence-based implementation of adult asthma guidelines in the emergency department: a controlled trial. *Emergency Medicine Australasia* 2007;**19**(1):31-8.

Dranitsaris 1995 *{published data only}*

Dranitsaris G, Warr D, Puodziunas A. A randomized trial of the effects of pharmacist intervention on the cost of antiemetic therapy with ondansetron. *Supportive Care in Cancer* 1995;**3**:183-9.

Dulko 2010 *{published data only}*

Dulko D, Hertz E, Julien J, Beck S, Mooney K. Implementation of cancer pain guidelines by acute care nurse practitioners using an audit and feedback strategy. *Journal of the American Academy of Nurse Practitioners* 2010;**22**:45-55.

Echouffo-Tcheugui 2009 *{published data only}*

Echouffo-Tcheugui JB, Simmons RK, Williams KM, Barling RS, Prevost AT, Kinmonth AL, et al. The ADDITION-Cambridge trial protocol: a cluster- randomised controlled trial of screening for type 2 diabetes and intensive treatment for screen-detected patients. *BMC Public Health* 2009;**9**:136.

Elnicki 1998 *{published data only}*

Elnicki DM, Layne RD, Ogden PE, Morris DK. Oral versus written feedback in medical clinic. *Journal of General Internal Medicine* 1998;**13**:155-8.

Everett *{published data only}*

Everett GD, deBlois CS, Chang PF, Holets T. Effect of cost education, cost audits, and faculty chart review on the use of laboratory services. *Archives of Internal Medicine* 1983;**143**:942-4.

Fallowfield 2002 *{published data only}*

Fallowfield L, Jenkins V, Farewell V, Saul J, Duffy A, Eves R. Efficacy of a Cancer Research UK communication skills training model for oncologists: a randomised controlled trial. *Lancet* 2002;**359**(9307):650-6.

Feder 1995 *{published data only}*

Feder G, Griffiths C, Highton C, Eldridge S, Spence M, Southgate L. Do clinical guidelines intorduced with practice based education improve care of asthmatic and dibetic patients? A randomised controlled trial in general practices in east London. *BMJ* 1995;**311**:1473-8.

Ferreira 2005 *{published data only}*

Ferreira MR, Dolan NC, Fitzgibbon ML, Davis TC, Gorby N, Ladewski L, et al. Health care provider-directed intervention to increase colorectal cancer screening among veterans: Results of a randomized controlled trial. *Journal of Clinical Oncology* 2005;**23**(7):1548-54.

Fick 2004 *{published data only}*

Fick DM, Maclean JR, Rodriguez NA, Short L, Heuvel RV, Waller JL, et al. A randomized study to decrease the use of potentially inappropriate medications among community-dwelling older adults in a southeastern managed care organization. *American Journal of Managed Care* 2004;**10**:761-8.

Fihn 2004 *{published data only}*

Fihn DS, McDonell MB, Diehr P, Anderson SM, Bradley KA, Au DH, et al. Effects of sustained audit/feedback on self-reported health status of primary care patients. *The American Journal of Medicine* 2004;**116**:241-8.

Finkelstein 2001 *{published data only}*

Finkelstein JA, Davis RL, Dowell SF, Metlay JP, Soumerai SB, Rifas-Shiman SL, et al. Reducing antibiotic use in children: a randomized trial in 12 practices. *Pediatrics* 2001;**108**(1):1-7.

Finkelstein 2005 *{published data only}*

Finkelstein JA, Lozano P, Fuhlbrigge AL, Carey VJ, Inui TS, Soumerai SB, et al. Practice-level effects of interventions to improve asthma care in primary care settings: The pediatric asthma care patient outcomes research team. *Health Services Research* 2005;**40**:1737-57.

Finkelstein 2008 *{published data only}*

Finkelstein JA, Huang SS, Kleinman K, Rifas-Shiman SL, Stille CJ, Daniel J, et al. Impact of a 16-community trial to promote judicious antibiotic use in Massachusetts. *Pediatrics* 2008;**121**(1):e15-23.

Frame 1994 *{published data only}*

Frame PS, Zimmer JG, Werth PL, Hall WJ, Eberly SW. Computer-based vs manual health maintenance tracking. A controlled trial. *Archives of Family Medicine* 1994;**3**:581-8.

Freeborn 1997 *{published data only}*

Freeborn DK, Shye D, Mullooly JP, Eraker S, Romeo J. Primary care physicians' use of lumbar spine imaging tests: effects of guidelines and practice pattern feedback. *Journal of General Internal Medicine* 1997;**12**:619-25.

Fretheim 2006 *{published data only}*

Fretheim A, Oxman AD, Håvelsrud K, Treweek S, Kristoffersen DT, Bjørndal A. Rational prescribing in primary care (RaPP): a cluster randomized trial of a tailored intervention. *PLoS Medicine* 2006;**3**(6):e134.

Furniss 2000 *{published data only}*

Furniss L, Burns A, Craig SKL, Scobie S, Cooke J, Farafher B. Effects of a pharmacists medication review in nursing homes. *British Journal of Psychiatry* 2000;**176**:563-7.

Ganz 2005 *{published data only}*

Ganz PA, Farmer MM, Belman MJ, Garcia CA, Streja L, Dietrich AJ, et al. Results of a randomized controlled trial to

increase colorectal cancer screening in a managed care health plan. *Cancer* 2005;**104**:2072-83.

Garrouste-Orgeas 2010 *{published data only}*

Garrouste-Orgeas, Tabah A, Schwebel C, Vesin A, Soufir L, Adrie C. A multifaceted program for improving quality of care in critically ill patients: The IATROREF Study (Part 3). Intensive Care Medicine. 2010:S89.

Gask 1991 *{published data only}*

Gask L, Goldberg D, Boardman J, Craig T, Goddard C, Jones O, et al. Training general practitioners to teach psychiatric interviewing skills: an evaluation of group training. *Medical Education* 1991;**25**(5):444-51. [MEDLINE: 92099912]

Gerbert 1988 *{published data only}*

Gerbert B, Maguire B, Badner V, Greenspan D, Greenspan J, Barnes D, et al. Changing dentists' knowledge, attitudes and behaviours relating to AIDS: a controlled educational intervention. *Journal of the American Dental Association* 1988;**116**:851-4.

Goderis 2010 *{published data only}*

Goderis G, Borgermans L, Grol R, Van Den Broeke C, Boland B, Verbeke G, et al. Start improving the quality of care for people with type 2 diabetes through a general practice support program: a cluster randomized trial. *Diabetes Research and Clinical Practice* 2010;**88**:56-64.

Goldberg 1980 *{published data only}*

Goldberg DP, Steele JJ, Smith C, Spivey L. Training family doctors to recognise psychiatric illness with increased accuracy. *Lancet* 1980;**6**:521-3.

Goldberg 1998 *{published data only}*

Goldberg HI, Wagner EH, Fihn SD, Martin DP, Horowitz CR, Christensen DB, et al. A randomized controlled trial of CQI teams and academic detailing: can they alter compliance with guidelines?. *Journal on Quality Improvement* 1998;**24**(3):130-42.

Grimshaw 1998 *{published data only}*

* Grimshaw J. Evaluation of four quality assurance initiatives to improve out-patient referrals from general practice to hospital. Thesis, University of Aberdeen 1998.

Gunn 2003 *{published data only}*

Gunn J, Southern D, Chondros P, Thomson P, Robertson K. Guidelines for assessing postnatal problems: introducing evidence-based guidelines in Australian general practice. *Family Practice* 2003;**20**:382-9.

Hall 2001 *{published data only}*

Hall L, Eccles M, Barton R, Steen N, Campbell M. Is untargeted outreach visiting in primary care effective? A pragmatic randomized controlled trial. *Journal of Public Health Medicine* 2001;**23**:109-13.

Hampshire 1999 *{published data only}*

Hampshire A, Blair M, Crown N, Avery A, Williams I. Action research: a useful method of promoting change in primary care?. *Family Practice* 1999;**16**(3):305-11.

Hanlon 1996 *{published data only}*

Hanlon JT, Weinberger M, Samsa GP, Kenneth E, Uttech KM, Lewis IK, et al. A randomized controlled trial of a clinical pharmacists intervention to improve inappropriate prescribing in elderly outpatients with polypharmacy. *The American Journal of Medicine* 1996;**100**(4):428-37.

Harbarth 2002 *{published data only}*

Harbarth S, Pittet D, Grady L, Zawacki A, Potter-Bynoe G, Samore MH, et al. Interventional study to evaluate the impact of an alcohol-based hand gel in improving hand hygiene compliance. *The Pediatric Infectious Disease Journal* 2002;**21**:489-95.

Harewood 2008 *{published data only}*

Harewood GC, Murray F, Winder S, Patchett S. Evaluation of formal feedback on endoscopic competence among trainees: the EFFECT trial. *Irish Journal of Medical Science* 2008;**177**(3):253-6.

Hargraves 1996 *{published data only}*

Hargraves JL, Palmer RH, Orav EJ, Wright EA. Practice characteristics and performance of primary care practitioners. *Medical Care* 1996;**34**(9):67-76.

Harris 2005 *{published data only}*

Harris SB, Leiter LA, Webster-Bogaert S, Van DM, O'Neill C. Teleconferenced educational detailing: diabetes education for primary care physicians. *Journal of Continuing Education in the Health Professions* 2005;**25**:87-97.

Hartlaub 1993 *{published data only}*

Hartlaub PP, Barrett PH, Marine WM, Murphy JR. Evaluation of an intervention to change benzodiazepine-prescribing behavior in a prepaid group practice setting. *American Journal of Preventive Medicine* 1993;**9**:346-52.

Henderson 1979 *{published data only}*

Henderson D, D´Alessandri R, Westfall B, Moore R, Smith R, Scobbo, et al. Hospital cost containment: a little knowledge helps. *Clinical Research* 1979;**27**:297A.

Hetlevik 1998 *{published data only}*

* Hetlevik I, Holmen J, Krüger Ø. Implementing clinical guidelines in the treatment of hypertension in general practice. *Scandinavian Journal of Primary Health Care* 1999;**17**:35-40.

Hetlevik I, Holmen J, Krüger Ø, Kristensen P, Iversen H. Implementing clinical guidelines in the treatment of hypertension in general practice. *Blood Pressure* 1998;**7**:270-6.

Hinchey 2010 *{published data only}*

Hinchey JA, Shephard T, Tonn ST, Ruthazer R, Hermann RC, Selker HP, et al. The Stroke Practice Improvement Network: a quasiexperimental trial of a multifaceted intervention to improve quality. *Journal of Stroke and Cerebrovascular Diseases* 2010;**19**:130-7.

Hirsch 2002 *{published data only}*

Hirsch IB, Goldberg HI, Ellsworth A, Evans TC, Herter CD, Ramsey SD, et al. A multifaceted intervention in support of

diabetes treatment guidelines: a cont trial. *Diabetes Research and Clinical Practice* 2002;**58**:27-36.

Hogg 2008 *{published data only}*

Hogg W, Lemelin J, Graham ID, Grimshaw J, Martin C, Moore L, et al. Improving prevention in primary care: evaluating the effectiveness of outreach facilitation. *Family Practice* 2008;**25**:40-8.

Holleman 1996 *{published data only}*

Holleman DR, .Simel DL. Effectiveness of automatic diagnostic test result feedback on outpatient laboratory and radiology testing in veterans. A controlled trial. *Medical Care* 1996;**34**:857-61.

Horbar 2004 *{published data only}*

Horbar JD, Carpenter JH, Buzas J, Soll RF, Suresh G, Bracken MB, et al. Collaborative quality improvement to promote evidence based surfactant for preterm infants: a cluster randomised trial. *BMJ* 2004;**329**:1004.

Horowitz 1996 *{published data only}*

Horowitz CR, Goldberg HI, Martin DP, Wagner EH, Fihn SD, Christensen DB, et al. Conducting a randomized controlled trial of CQI and academic detailing to implement clinical guidelines. *Journal on Quality Improvement* 1996;**22**(11):734-50.

Howe 1996 *{published data only}*

Howe A. Detecting psychological distress: can general practitioners improve their performance?. *British Journal of General Practice* 1996;**46**:407-10.

Hulscher 1997 *{published data only}*

Hulscher ME, Van Drenth BB, Van der Wouden JC, Mokkink HG, Van Weel C, Grol RP. Changing preventive practice: a controlled trial on the effects of outreach visits to organise prevention of cardiovascular disease. *Quality in Health Care* 1997;**6**:19-24.

Ilag 2003 *{published data only}*

Ilag LL, Martin CL, Tabaei BP, Isaman DJ, Burke R, Greene DA, et al. Improving diabetes processes of care in managed care. *Diabetes Care* 2003;**26**:2722-7.

Jaen 2010 *{published data only}*

Jaen CR, Crabtree BF, Palmer RF, Ferrer RL, Nutting PA, Miller WL, et al. Methods for evaluating practice change toward a patient-centered medical home. *Annals of Family Medicine* 2010;**8**(Suppl 1):S9-20.

Jans 2001 *{published data only}*

Jans MP, Schellevis FG, Le Coq EM, Bezemer PD, Van Eijk JT. Health outcomes of asthma and COPD patients: the evaluation of a project to implement guidelines in general practice. *International Journal for Quality in Health Care* 2001;**13**:17-25.

Johansen 1997 *{published data only}*

Johansen A. Using audit to improve senior officer training. *Postgraduate Medical Journal* 1997;**73**:798-801.

Johnson 1976 *{published data only}*

Johnson RE, Campbell WH, Azevedo DJ, Christensen DB. Studying the impact of patient drug profile in an HMO. *Medical Care* 1976;**15**:799-807.

Jones 1996 *{published data only}*

Jones HE, Cleave B, Zinman B, Szalai JP, Nichol HL, Hoffman BR. Efficacy of feedback from quarterly laboratory comparison in maintaining quality of a hospital capillary blood glucose monitoring program. *Diabetes Care* 1996;**19**(2):168-70.

Kafuko 1999 *{published data only}*

Kafuko JM, Zirabamuzaale, Bagena D. Rational drug use in rural health units of Uganda:effect of national standard treatment guidelines on rational drug use. 1st International Conference on Improving Use of Medications. 1999.

Katz 2004 *{published data only}*

Katz DA, Muehlenbruch DR, Brown RL, Fiore MC, Baker TB. Effectiveness of implementing the agency for healthcare research and quality smoking cessation clinical practice guideline: A randomized, controlled trial.. *Journal of the National Cancer Institute* 2004;**96**:594-603.

Kerse 1999 *{published data only}*

Kerse NM, Flicker L, Jolley D, Arroll B, Young D. Improving the health behaviours of elderly people: randomised controlled trial of a general practice education programme. *BMJ* 1999;**319**:683-7.

Kinney 2003 *{published data only}*

Kinney ED, Kennedy J, Cook CA, Freedman JA, Lane KA, Hui SL. A randomized trial of two quality improvement strategies implemented in a statewide public community-based, long-term care program. *Medical Care* 2003;**41**:1048-57.

Kirwin 2010 *{published data only}*

Kirwin JL, Cunningham RJ, Sequist TD. Pharmacist recommendations to improve the quality of diabetes care: a randomized controlled trial. *Journal of Managed Care Pharmacy* 2010;**16**:104-13.

Kroenke 1990 *{published data only}*

Kroenke K, Pinholt EM. Reducing polypharmacy in the elderly. *Journal of American Geriatrics Society* 1990;**38**:31-6.

Kuilboer 2006 *{published data only}*

Kuilboer MM, Van Wijk MAM, Mosseveld M, Van der Does E, De Jongste JC, Overbeek SE, et al. Computed critiquing integrated into daily clinical practice affects physicians' behavior--a randomized clinical trial with AsthmaCritic. *Methods of information in medicine* 2006;**45**:447-54.

Labarere 2007 *{published data only}*

Labarere J, Bosson JL, Sevestre MA, Sellier E, Richaud C, Legagneux A. Intervention targeted at nurses to improve venous thromboprophylaxis. *International Journal for Quality in Health Care* 2007;**19**:301-8.

Lafata 2007 *{published data only}*

Lafata JE, Gunter MJ, Hsu J, Kaatz S, Krajenta R, Platt R, e al. Academic detailing to improve laboratory testing among outpatient medication users. *Medical Care* 2007;**45**(10):966-72.

Lassen 1992 *{published data only}*

Lassen LC, Kristensen FB. Peer comparison feedback to achieve rational and economical drug therapy in general practice: a controlled intervention study. *Scandinavian Journal of Primary Health Care* 1992;**10**:76-80.

Lemelin 2001 *{published data only}*

Lemelin J, Hogg W, Baskerville N. Evidence to action: a tailored multifaceted approach to changing family physician practice patterns and improving preventive care. *Scandinavian Journal of Primary Health Care* 2001;**164**:757-63.

Lenderink 2010 *{published data only}*

Lenderink AF, Spreeuwers D, Van der Klink JJL, Van Dijk FJH. Information and feedback to improve occupational physicians' reporting of occupational diseases: a randomised controlled trial. *International Archives of Occupational and Environmental Health* 2010;**83**:381-8.

Leviton 1999 *{published data only}*

Leviton LC, Goldenberg RL, Baker CS, Schwartz RM, Freda MC, Fish LJ, et al. Methods to encourage the use of antenatal corticosteroid therapy for fetal maturation. *JAMA* 1999;**281**(1):46-52.

Linn 1980 *{published data only}*

Linn LS, Yager J. The effect of screening, sensitization, and feedback on notation of depression. *Journal of Medical Education* 1980;**55**:942-9.

Luders 2010 *{published data only}*

Luders S, Schrader J, Schmieder RE, Smolka W, Wegscheider K, Bestehorn K. Improvement of hypertension management by structured physician education and feedback system: cluster randomized trial. *European Journal of Cardiovascular Prevention & Rehabilitation* 2010;**17**:271-9.

Lundborg 1999 *{published data only}*

Lundborg CS, Wahlström R, Oke T, Tomson G, Diwan VK. Influencing prescribing for urinary tract infection and asthma in primary care in Sweden: a randomized controlled trial of an interactive educational intervention. *Journal of Clinical Epidemiology* 1999;**52**:801-12.

MacCosbe 1985 *{published data only}*

MacCosbe PE, Gartenberg G. Modifying empiric antibiotic prescribing: Experience with one strategy in a medical residency program. *Hospital Formulary* 1985;**20**:986-99.

MacGowan 1996 *{published data only}*

MacGowan AP, Feeney R, Brown I, Mcculloch S, Reeves D, Lovering A. Routine feedback to GPs who request microbiological tests is effective. *BMJ* 1996;**312**:1481.

Madridejos-Mora 2004 *{published data only}*

Madridejos-Mora R, Amado-Guirado E, Pérez-Rodríguez MT. Effectiveness of the combination of feedback and educational recommendations for improving drug prescription in general practice. *Medical Care* 2004;**42**:643-8.

Mandel 1985 *{published data only}*

Mandel I, Franks P, Dickinson J. Improving physician compliance with preventive medicine guidelines. *The Journal of Family Practice* 1985;**21**(3):223-4. [MEDLINE: 85291598]

Manfredi 1998 *{published data only}*

Manfredi C, Czaja R, Freels S, Trubitt M, Warnecke R, Lacey L. Improving cancer screening in physcians practices serving low-income and minority populations. *Archives of Family Medicine* 1998;**7**:329-37.

Manheim 1990 *{published data only}*

Manheim LM, Feinglass J, Hughes R, Martin GJ, Conrad K, Hughes EF. Training house officers to be cost conscious. Effects of an educational intervention on charges and length of stay. *Medical Care* 1990;**28**:29-42.

Manning 1986 *{published data only}*

Manning PR, Lee PV, Clintworth WA, Denson TA, Oppenheimer PR, Gilman NJ. Changing prescribing practices through individual continuing education. *JAMA* 1986;**256**:230-2.

Martin 2007 *{published data only}*

Martin WE, Miller SC, Welch LC, Burrill J. Improving access to hospice: the Physician Feedback and Reminders to Improve Access to Hospice (PFRIAH) study. *Medicine and Health, Rhode Island* 2007;**90**(12):388-90.

Mayefsky 1993 *{published data only}*

Mayefsky JH, Foye HR. Use of a chart audit: teaching well child care to paediatric house officers. *Medical Education* 1993;**27**:170-4.

Mazzuca 1988 *{published data only}*

Mazzuca SA, Vinicor F, Cohen SJ, Norton JA, Fineberg NS, Fineberg SE. The Diabetes Education study: a controlled trial of the effects of intensive instructions of internal medicine residents on the management of diabetes mellitus. *Journal of General Internal Medicine* 1988;**3**:1-8.

McDermott 2003 *{published data only}*

McDermott R, Tulip F, Schmidt B, Sinha A. Sustaining better diabetes care in remote indigenous Australian communities. *BMJ* 2003;**327**:428-30.

McDonel 1997 *{published data only}*

McDonel EC, Bond GR, Salyers M, Fekete D, Chen A, McGrew JH, et al. Implementing assertive community treatment programs in rural settings. *Administration and Policy in Mental Health* 1997;**25**(2):153-73.

McPhee 1989 *{published data only}*

* Fordham D, McPhee SJ, Bird JA, Rodnick JE, Detmer WM. The cancer prevention reminder system. *Clinical Computing* 1990;**7**(5):289-95.

McPhee SJ, Bird JA, Jenkins CNH, Fordham D. Promoting cancer screening. A randomized controlled trial of three interventions. *Archives of Internal Medicine* 1989;**149**:1868-72.

Meehan 2001 *{published data only}*

Meehan TP, Weingarten SR, Holmboe ES, Mathur D, Wang Y, Petrillo MK, et al. A statewide initiative to improve the care of hospitalized pneumonia patients: The Connecticut Pneumonia Pathway Project. *American Journal of Medicine* 2001;**111**:203-10.

Mertz 2010 *{published data only}*

Mertz D, Dafoe N, Walter SD, Brazil K, Loeb M. Effect of a multifaceted intervention on adherence to hand hygiene among healthcare workers: A cluster-randomized trial. *Infection Control and Hospital Epidemiology* 2010;**31**:1170-6.

Metlay 2007 *{published data only}*

Metlay JP, Camargo CAJ, MacKenzie T, McCulloch C, Maselli J, Levin SK, IMPAACT Investigators. Cluster-randomized trial to improve antibiotic use for adults with acute respiratory infections treated in emergency departments. *Annals of Emergency Medicine* 2007;**50**(3):221-30.

Meyer 1991 *{published data only}*

Meyer TJ, Van Kooten D, Marsh S, Prochazka AV. Reduction of polypharmacy by feedback to clinicians. *Journal of General Internal Medicine* 1991;**6**:133-6.

Moongtui 2000 *{published data only}*

Moongtui W, Gauthier DK, Turner JG. Using peer feedback to improve handwashing and glove usage among Thai health care workers. *American Journal of Infection Control* 2000;**28**:365-9.

Mourad 2010 *{published data only}*

Mourad S, Hermens RPMG, Nelen WLDM, Grol RPTM, Kremer JAM. Human Reproduction. 2010 Conference:i258.

Munroe 1997 *{published data only}*

Munroe WP, Kunz K, Dalmandy-Israel C, Potter L, Schonfeld WH. Economic evaluation of pharmacist involving in disease management in a community pharmacy setting. *Clinical Therapeutics* 1997;**19**(1):113-23.

Myers 2004 *{published data only}*

Myers RE, Turner B, Weinberg D, Hyslop T, Hauck WW, Brigham T, et al. Impact of a physician-oriented intervention on follow-up in colorectal cancer screening. *Preventive Medicine* 2004;**38**:375-81.

Nattinger 1989 *{published data only}*

Nattinger AB, Panzer RJ, Janus J. Improving the utilization of screening mammography in primary care practices. *Archives of Internal Medicine* 1989;**149**(9):2087-92. [MEDLINE: 89373341]

Nicolas 1996 *{published data only}*

Nicolas F, Raimondeau J, Blanloeil Y, Le Conte P, Villers D, Touze MD. Cost analysis, Consensus Conference, medical audit. *Journal D'Economie Medicale* 1996;**14**:145-57.

North of England1992 *{published data only}*

North of England Study of Standards and Performance in General Practice. Medical audit in general practice. I: Effects on doctors' clinical behaviour for common childhood conditions. *BMJ* 1992;**304**(6840):1480-4. [MEDLINE: 92305733]

North of England Study of Standards and Performance in General Practice. Medical audit in general practice. II: Effects on health of patients with common childhood conditions. *BMJ* 1992;**304**(6840):1484-8. [MEDLINE: 92305734]

Nyman 1995 *{published data only}*

Nyman JA, Akhtar MR, Feldman R. Does publishing the parameters that trigger review of Medicare claims change provider behavior? Results of the parameter release study. *Medical Care* 1995;**33**:1022-34.

O'Connor 1996 *{published data only}*

O'Connor GT, Plume SK, Olmstead EM, Morton JR, Maloney CT, Nugent WC, et al. A regional intervention to improve the hospital mortality associated with coronary artery bypass graft surgery. The Northern New England Cardiovascular Disease Study Group. *JAMA* 1996;**275**:841-6.

O'Connor 2005 *{published data only}*

O'Connor PJ, Desai J, Solberg LI, Reger LA, Crain AL, Asche SE, et al. Randomized trial of quality improvement intervention to improve diabetes care in primary care settings. *Diabetes Care* 2005;**28**:1890-7.

Ogwal-Okeng 2001 *{published data only}*

Ogwal-Okeng JW, Anokbonggo WW, Birungi H. Prescribing audit with feedback intervention in six regional hospitals and Mulagi referral/teaching hospitals. International conference on Improving use of Medicines. http:/www.who.int/dap-icium/posters/2C3. 2001.

Ornstein 2010 *{published data only}*

Ornstein S, Nemeth LS, Jenkins RG, Nietert PJ. Colorectal cancer screening in primary care: translating research into practice. *Medical Care* 2010;**48**:900-6.

Ottolini 1998 *{published data only}*

Ottolini MC, Greenberg L. Development and evaluation of a CD-ROM computer program to teach residents telephone management. *Pediatrics* 1998;**101**(3):E2.

Overbeek 2010 *{published data only}*

Overbeek LI, Hermens RP, Van Krieken JH, Adang EM, Casparie M, Nagengast FM, et al. Electronic reminders for pathologists promote recognition of patients at risk for Lynch syndrome: Cluster-randomised controlled trial. *Virchows Arch* 2010;**456**:653-9.

Papa 1999 *{published data only}*

Papa FJ, Aldrich D, Schumacker RE. The effects of immediate online feedback upon diagnostic performance. *Academic Medicine* 1999;**74**:S16-S18.

Patel 2010 *{published data only}*

Patel L, Glasscoe C, Dixon C, Dyer K, Southern KW. Pediatric Pulmonology. 2010 Conference:441.

Payne 1978 *{published data only}*

Payne BC, Lyons TF, Neuhaus E, Kolton M, Dwarshius L. Method of Evaluating and Improving Ambulatory Medical Care. Final Report and Executive Summary. Final report 30. NTIS 1978; Vol. Order Number: PB80 178585.

Pearson 2001 *{published data only}*

Pearson SD, Kleefield SF, Soukop JR, Cook EF, Lee TH. Critical pathways intervention to reduce length of hospital stay. *The American Journal of Medicine* 2001;**110**:175-80.

Performance 2006 *{published data only}*

Using performance feedback to change provider behaviour. Performance improvement advisor 2006; Vol. 10, issue 5:58-9.

Peters-Klimm 2008 *{published data only}*

Peters-Klimm F, Muller-Tasch T, Remppis A, Szecsenyi J, Schellberg D. Improved guideline adherence to pharmacotherapy of chronic systolic heart failure in general practice--results from a cluster-randomized controlled trial of implementation of a clinical practice guideline. *Journal of Evaluation in Clinical Practice* 2008;**14**(5):823-9.

Peters-Klimm 2009 *{published data only}*

Peters-Klimm F, Campbell S, Muller-Tasch T, Schellberg D, Gelbrich G, Herzog W, et al. Primary care-based multifaceted, interdisciplinary medical educational intervention for patients with systolic heart failure: lessons learned from a cluster randomised controlled trial. *Trials [Electronic Resource]* 2009;**10**:68.

Pfirter 2010 *{published data only}*

Pfirter G, Lapertosa S, Villagra M, Caporale JE, Gonzalez C, Clark C, et al. Diabetologia. 2010 Conference:S109.

Pit 2007 *{published data only}*

Pit SW, Byles JE, Henry DA, Holt L, Hansen V, Bowman DA. A Quality Use of Medicines program for general practitioners and older people: a cluster randomised controlled trial. *Medical Journal of Australia* 2007;**187**(1):23-30.

Pugh 1989 *{published data only}*

Pugh JA, Frazier LM, DeLong E, Wallace AG, Ellenbogen P, Linfors E. Effect of daily charge feedback on inpatient charges and physician knowledge and behavior. *Archives of Internal Medicine* 1989;**149**:426-9.

Putnam 1985 *{published data only}*

Putnam RW, Curry L. Impact of patient care appraisal on physicians behaviour in the office setting. *Canadian Medical Association Journal* 1985;**132**:1025-9.

Putnam W, Curry L. Primary care appraisal in the ambulatory setting: effectiveness of a continuing medical education tool. 419 RCT-C in register.

Quilitch 1975 *{published data only}*

Quilitch HR. A comparison of three staff-management procedures. *Journal of Applied Behavior Analysis* 1975;**8**:59-66.

Raisch 1999 *{published data only}*

Raisch DW, .Sleath BL. Using feedback letters to influence the use of antiulcer agents in a Medicaid program. *Journal of General Internal Medicine* 1999;**14**:145-50.

Rascati 1996 *{published data only}*

Rascati KL, Okano GJ, Burch C. Evaluation of physician intervention letters. *Medical Care* 1996;**34**:760-6.

Reid 1977 *{published data only}*

Reid RA, Lantz KH. Physician profiles in training the graduate internist. *Journal of Medical Education* 1977;**52**:300-5.

Restuccia 1982 *{published data only}*

Restuccia JD. The effect of concurrent feedback in reducing inappropriate hospital utilization. *Medical Care* 1982;**20**(1):46-62. [MEDLINE: 82194054]

Reuther 2010 *{published data only}*

Reuther L, Schultz-Larsen P, Damsgaard J, Gilsa HD, Munck A, Sondergaard J, et al. Basic & Clinical Pharmacology & Toxicology Conference. 2010:542.

Rhew 1999 *{published data only}*

Rhew DC, Glassman PA, Goetz MB. Improving pneumococcal vaccine rates. Nurse protocols versus clinical reminders. *Journal of General Internal Medicine* 1999;**14**:351-6.

Rollman 2002 *{published data only}*

Rollman BL, Hanusa BH, Lowe HJ, Gilbert T, Kapoor WN, Schulberg HC. A randomized trial using computerized decision support to improve treatment of major depression in primary care. *Journal of General internal Medicine* 2002;**17**:493-503.

Roski 1998 *{published data only}*

Roski J. Changing practice patterns as a result of implementing the Agency for Health Care Policy and Research guidelines in 20 primary care clinics. *Tobacco Control* 1998;**7 Suppl**:S19-20; discussion S24-5.

Rubenstein 1989 *{published data only}*

Rubenstein LV, Calkin DR, Young RT, Cleary PD, Fink A, Kosecoff J, et al. Improving patient function: a randomized trial of functional disability screening. *Annals of Internal Medicine* 1989;**111**(10):836-42.

Rubenstein 1999 *{published data only}*

Rubenstein LV, Jackson-Tricke M, Unützer J, Miranda J, Minnium K, Pearson ML, et al. Evidence-based care for depression in managed primary care practices. *Health Affairs* 1999;**18**(5):89-105.

Sanazaro 1978 *{published data only}*

Sanazaro PJ, Worth RM. Concurrent quality assurance in hospital care. Report of a study by Private Initiative in PSRO. *The New England Journal of Medicine* 1978;**298**:1171-7.

Seers 2004 *{published data only}*

Seers K, Crichton N, Carroll D, Richards S, Saunders T. Evidence-based postoperative pain management in nursing: is a randomized-controlled trial the most appropriate design?. *Journal of Nursing Management* 2004;**12**:183-93.

Shaughnessy 1991 *{published data only}*

Shaughnessy AF, DAmico F, Nickel RO. Improving prescription-writing skills in a family practice residency. *DICP, The Annals of Pharmacotherapy* 1991;**25**:17-21.

Simon 2000 *{published data only}*

Simon GE, VonKorff M, Rutter C, Wagner E. Randomised trial in monitoring, feedback, and management of care by telephone to improve treatment of depression in primary care. *BMJ* 2000;**320**:550-4.

Simunovic 2010 *{published data only}*

Simunovic M, Coates A, Goldsmith CH, Thabane L, Reeson D, Smith A, et al. The cluster-randomized Quality Initiative in Rectal Cancer trial: evaluating a quality-improvement strategy in surgery. *Canadian Medical Association Journal* 2010;**182**:1301-6.

Smeele 1998 *{published data only}*

Smeele IJM, Grol RPTM, Van Schayck CP. Implementation of CARA standards in general practice: a randomised camparative study into the effectivity of a monitoring-feedback system and advancement of expertise on general practitioners. *Nederlands Tijdschrift voor Geneeskunde* 1998;**142**:1422.

Smith 1995 *{published data only}*

Smith D, Christensen DB, Stergachis A, Holmes G. A randomized controlled trial of a drug use review intervention for sedative hypnotic medications. *Prenatal Diagnosis* 1998;**15**:1013-21.

Spector 1989 *{published data only}*

Spector WD, Drugovich ML. Reforming nursing home quality regulation. Impact on cited deficiencies and nursing home outcomes. *Medical Care* 1989;**27**(8):789-801. [MEDLINE: 89329844]

Steele 1989 *{published data only}*

Steele MA, Bess DT, Franse VL, Graber SE. Cost effectiveness of two interventions for reducing outpatient prescribing costs. *DICP: The Annals of Pharmacotherapy* 1989;**23**(6):497-500.

Stewart 2005 *{published data only}*

Stewart M, Marshall JN, Stbye T, Feightner JW, Brown JB, Harris S, et al. Effectiveness of case-based on-line learning of evidence-based practice guidelines. *Family Medicine* 2005;**37**:131-8.

Strasser 2008 *{published data only}*

Strasser DC, Falconer JA, Stevens AB, Uomoto JM, Herrin J, Bowen SE, et al. Team training and stroke rehabilitation outcomes: a cluster randomized trial. *Archives of Physical Medicine and Rehabilitation* 2008;**89**(1):10-5.

Strikwerda 1994 *{published data only}*

Strikwerda P, Bootsma-de Langen AM, Berghuis F, Meyboom-de Jong B. Drug therapy in a nursing home; favorable effect of feedback by the pharmacist on family physician's prescribing behavior. *Nederlands Tijdschrift voor Geneeskunde* 1994;**138**:1770-4.

Sundaram 2009 *{published data only}*

Sundaram V, Lazzeroni LC, Douglass LR, Sanders GD, Tempio P, Owens DK. A randomized trial of computer-based reminders and audit and feedback to improve HIV screening in a primary care setting. *International journal of STD & AIDS* 2009;**20**:527-33.

Szczepura 1994 *{published data only}*

Szczepura A, Wilmot J, Davies C, Fletcher J. Effectiveness and cost of different strategies for information feedback in general practice. *The British Journal of General Practice* 1994;**44**(378):19-24. [MEDLINE: 94145722]

Taylor 1997 *{published data only}*

Taylor EA, Thomas G, Cantrill JA. Changes in prescribing following a pharmicists-led audit of ulcer-healing therapy in general practice. *The Pharmaceutical Journal* 1997;**259**:R6.

The SUPPORT 1995 *{published data only}*

The SUPPORT Principal Investigators. A controlled trial to improve care for seriously ill hospitalized patients. *JAMA* 1995;**274**:1591-8.

Thompson 2000 *{published data only}*

Thompson RS, Rivara FP, Thompson DC, Barlow WE, Sugg NK, Maiuro RD, et al. Identification and management of domestic violence: a randomized trial. *American Journal of Preventive Medicine* 2000;**19**(4):253-63.

Van Bruggen 2008 *{published data only}*

Van Bruggen R, Gorter KJ, Stolk RP, Verhoeven RP, Rutten GEHM. Implementation of locally adapted guidelines on type 2 diabetes. *Family Practice* 2008;**25**(6):430-7.

Van der Sanden 2005 *{published data only}*

Van der Sanden WJM, Mettes DG, Plasschaert AJM, Grol RPTM, Mulder J, Verdonschot EH. Effectiveness of clinical practice guideline implementation on lower third molar management in improving clinical decision-making: a randomized controlled trial. *European Journal of Oral Sciences* 2005;**113**:349-54.

Van der Weijden 1998 *{published data only}*

Van der Weijden T, Grol RPTM, Schouten BJ, Knottnerus JA. Barriers to working according to cholesterol guidelines: A randomized controlled trial on implementation of national guidelines in 20 general practices. *European Journal of Public Health* 1998;**8**:113-8.

Velikova 2004 *{published data only}*

Velikova G, Booth L, Smith AB, Brown PM, Lynch P, Brown JM, et al. Measuring quality of life in routine oncology practice improves communication and patient well-being: a randomized controlled trial. *Journal of Clinical Oncology* 2004;**22**:714-24.

Verstappen 2004 b *{published data only}*

Verstappen WHJM, Van Merode F, Grimshaw J, Dubois WI, Grol RPTM, Van der Weijden T. Comparing cost effects of two quality strategies to improve test ordering in primary care: a randomized trial. *International Journal for Quality in Health Care* 2004;**16**:391-8.

Vinicor 1987 *{published data only}*

Vinicor F, Cohen SJ, Mazzuca SA, Moorman N, Wheeler M, Kuebler T, et al. DIABEDS: a randomized trial of the effects of physician and/or patient education on diabetes patient outcomes. *Journal of Chronic Diseases* 1987;**40**:345-56.

Walsh 2007 *{published data only}*

Walsh M, Laptook A, Kazzi SN, Engle WA, Yao Q, Rasmussen M, et al. A cluster-randomized trial of benchmarking and multimodal quality improvement to improve rates of survival free of bronchopulmonary dysplasia for infants with birth weights of less than 1250 grams. *Pediatrics* 2007;**119**:876-90.

Watkins 1981 *{published data only}*

Watkins CJ. Medical audit in general practice - fact or fantasy?. *The Journal of the Royal College of General Practitioners* 1981;**31**:141-5.

Weingarten 2000 *{published data only}*

Weingarten SR, Kim CS, Stone EG, Kristopaitis RJ, Pelter M, Sandhu M. Can peer-comparison feedback improve patient functional status?. *The American Journal of Managed Care* 2000;**6**:35-9.

Wells 2000 *{published data only}*

Wells KB, Sherbourne C, Schoenbaum M, Duan N, Meredith L, Unutzer J, et al. Impact of disseminating quality improvement programs for depression in managed primary care. *JAMA* 2000;**283**(2):212-20.

Welschen 2004 *{published data only}*

Welschen I, Kuyvenhoven MM, Hoes AW, Verheij TJ. Effectiveness of a multiple intervention to reduce antibiotic prescribing for respiratory tract symptoms in primary care: randomised controlled trial. *BMJ* 2004;**329**:431.

White 1995 *{published data only}*

White P, Atherton A, Hewett G, Howells K. Using information from asthma patients: a trial of information feedback in primary care. *BMJ* 1995;**311**:1065-9.

Wing 1987 *{published data only}*

Wing D, Duff HJ. Impact of a therapeutic drug monitoring program for digoxin. *Archives of Internal Medicine* 1987;**147**:1405-8.

Wing 1987 (II) *{published data only}*

Wing DS, Duff HJ. Evaluation of a therapeutic drug monitoring program frotheophylline in a teaching hospital. *Drug Intelligence and Clinical Pharmacy* 1987;**21**:702-5.

Winickoff 1985 *{published data only}*

Winickoff RN, Wilner S, Neisuler R, Barnett GO. Limitations of provider interventions in hypertension quality assurance. *American Journal of Public Health* 1985;**75**:43-6.

Winkens 1992 *{published data only}*

Winkens RA, Pop P, Grol RP, Kester AD, Knottnerus JA. Effect of feedback on test ordering behaviour of general practitioners. *BMJ* 1992;**304**:1093-6.

Winkens 1997 *{published data only}*

Winkens RAG, Knottnerus JA, Kester ADM, Grol RPTM, Pop P. Fitting a routine health-care activity into a randomized trial: an experiment possible without informed consent?. *Clinical Epidemiology* 1997;**50**(4):435-9.

Yano 2008 *{published data only}*

Yano EM, Rubenstein LV, Farmer MM, Chernof BA, Mittman BS, Lanto AB, et al. Targeting primary care referrals to smoking cessation clinics does not improve quit rates: Implementing evidence-based interventions into practice. *Health Services Research* 2008;**43**(5 P1):1637-61.

Young 2002 *{published data only}*

Young JM, DEste C, Ward JE. Improving family physicians' Use of evidence-based smoking cessation strategies: A cluster randomization trial. *Preventive Medicine* 2002;**35**(6):572-83.

Young JM, Ward JE. Randomised trial of intensive academic detailing to promote opportunistic recruitment of women to cervical screening by general practitioners.. *Australian and New Zealand Journal of Public Health* 2003;**27**:273-81.

Zermansky 2002 *{published data only}*

Zermansky AG, Petty DR, Raynor DK, Lowe CJ, Freemantle N, Vail A. Clinical medication review by a pharmacist of patients on repeat prescriptions in general practice: a randomised controlled trial. *Health Technology Assessment* 2002;**6**:1-86.

Zoutman 2010 *{published data only}*

Zoutman D, Douglas F. American Journal of Infection Control. 2010 Conference:E14-E15.

References to studies awaiting assessment

Bond 2011 *{published data only}*

Bond TC, Patel PR, Krisher J, Sauls L, Deane J, Strott K, et al. A Group-Randomized Evaluation of a Quality Improvement Intervention to Improve Influenza Vaccination Rates in Dialysis Centers. *American Journal of Kidney Diseases* 2011;**57**:283-90.

Daley 2011 *{published data only}*

Daley M, Shepard DS, Tompkins C, Dunigan R, Reif S, Perloff J, et al. Randomized Trial of Enhanced Profiling in Substance Abuse Treatment. *Administration and Policy in Mental Health and Mental Health Services Research* 2011;**38**:96-104.

Guldberg 2011 *{published data only}*

Guldberg TL, Vedsted P, Kristensen JK, Lauritzen T. Improved quality of Type 2 diabetes care following electronic feedback of

treatment status to general practitioners: a cluster randomized controlled trial. *Diabetic Medicine* 2011;**28**:325-32.

Ivers 2010 *{published data only}*

Ivers NM, Tu K, Francis J, Barnsley J, Shah B, Upshur R, et al. Feedback GAP: study protocol for a cluster-randomized trial of goal setting and action plans to increase the effectiveness of audit and feedback interventions in primary care. *Implementation Science* 2010;**5**:98.

LaPointe 2006 *{published data only}*

LaPointe NMA, Delong ER, Chen A, Hammill BG, Muhlbaier LH, Califf RM, et al. Multifaceted intervention to promote beta-blocker use in heart failure. *American Heart Journal* 2006;**151**:992-8.

Lopez-Picazo 2011 *{published data only}*

Lopez-Picazo JJ, Ruiz JC, Sanchez JF, Ariza A, Aguilera B. A randomized trial of the effectiveness and efficiency of interventions to reduce potential drug interactions in primary care. *American Journal of Medical Quality* 2011;**26**:145-53.

Mourad 2011 *{published data only}*

Mourad SM, Hermens RPMG, Liefers J, Akkermans RP, Zielhuis GA, Adang E, et al. A multi-faceted strategy to improve the use of national fertility guidelines; a cluster-randomized controlled trial. *Human Reproduction* 2011;**26**:817-26.

Palmer 1996 *{published data only}*

Palmer RH, Louis TA, Peterson HF, Rothrock JK, Strain R, Wright EA. What makes quality assurance effective? Results from a randomized, controlled trial in 16 primary care group practices. *Medical Care* 1996;**34**:SS29-SS39.

Sequist 2010 *{published data only}*

Sequist TD, Fitzmaurice GM, Marshall R, Shaykevich S, Marston A, Safran DG, et al. Cultural competency training and performance reports to improve diabetes care for black patients: a cluster randomized, controlled trial.. *Annals of Internal Medicine* 2010;**152**(1):40-6.

Smeets 2010 *{published data only}*

Smeets HM, De Wit NJ, Zuithoff NPA, Van Dijk PCM, Van der Lee APM, Hoes AW. A health insurance company-initiated multifaceted intervention for optimizing acid-suppressing drug prescriptions in primary care: a randomized controlled trial. *Archives of Internal Medicine* 2010;**170**:1264-70.

Williams 2011 *{published data only}*

Williams JB, Delong ER, Peterson ED, Dokholyan RS, Ou FS, Ferguson TB. Secondary prevention after coronary artery bypass graft surgery: findings of a national randomized controlled trial and sustained society-led incorporation into practice. *Circulation* 2011;**123**(1):39-45.

Additional references

Axt-Adam 1993

Axt-Adam P, Van der Wouden JC, Van der Does E. Influencing behavior of physicians ordering laboratory tests: a literature study. *Medical Care* 1993;**31**:784-94.

Balas 1996

Balas EA, Boren SA, Brown GD, Ewigman BG, Mitchell JA, Perkoff GT. Effect of physician profiling on utilization. Meta-analysis of randomized clinical trials. *Journal of General Internal Medicine* 1996;**11**:584-90.

BMJ 1992

[No authors listed]. Medical audit in general practice. I: Effects on doctors' clinical behaviour for common childhood conditions. North of England Study of Standards and Performance in General Practice. *BMJ* 1992;**304**(6840):1480-4.

Carlsen 2007

Carlsen B, Glenton C, Pope C. Thou shalt versus thou shalt not: a meta-synthesis of GPs' attitudes to clinical practice guidelines. *British Journal of General Practice* 2007 December;**57**(545):971-8.

Carver 1982

Carver CS, Scheier MF. Control theory: a useful conceptual framework for personality-social, clinical, and health psychology. *Psychological Bulletin* 1982 July;**92**(1):111-35.

Davidoff 2009

Davidoff F, Batalden P, Stevens D, Ogrinc G, Mooney SE, SQUIRE development group. Publication guidelines for quality improvement studies in health care: evolution of the SQUIRE project. *BMJ* 2009;**338**:a3152.

Davis 2006

Davis DA, Mazmanian PE, Fordis M, Van Harrison R, Thorpe KE, Perrier L. Accuracy of physician self-assessment compared with observed measures of competence: a systematic review. *JAMA* 2006 Sept;**296**(9):1094-102.

EPOC 2002

Cochrane Effective Practice and Organisation of Care Group (EPOC). Cochrane Effective Practice and Organisation of Care Review Group Data Collection Checklist. http://epoc.cochrane.org/sites/epoc.cochrane.org/files/uploads/datacollectionchecklist.pdf [Revised June 2002].

Farmer 2008

Farmer AP, Légaré F, Turcot L, Grimshaw J, Harvey E, McGowan JL, et al. Printed educational materials: effects on professional practice and health care outcomes. *Cochrane Database of Systematic Reviews* 2008, Issue 3. [DOI: 10.1002/14651858.CD004398.pub2]

Flodgren 2011

Flodgren G, Parmelli E, Doumit G, Gattellari M, O'Brien MA, Grimshaw J, et al. Local opinion leaders: effects on professional practice and health care outcomes. *Cochrane Database of Systematic Reviews* 2011, Issue 8. [DOI: 10.1002/14651858.CD000125.pub4]

Forsetlund 2009

Forsetlund L, Bjørndal A, Rashidian A, Jamtvedt G, O'Brien MA, Wolf F, et al. Continuing education meetings and workshops: effects on professional practice and health care outcomes.

Cochrane Database of Systematic Reviews 2009, Issue 2. [DOI: 10.1002/14651858.CD003030.pub2]

Foy 2002

Foy R, MacLennan G, Grimshaw J, Penney G, Campbell M, Grol R. Attributes of clinical recommendations that influence change in practice following audit and feedback. *Journal of Clinical Epidemiology* 2002;**55**:717-22.

Foy 2005

Foy R, Eccles MP, Jamtvedt G, Young J, Grimshaw JM, Baker R. What do we know about how to do audit and feedback? Pitfalls in applying evidence from a systematic review. *BMC Health Services Research* 2005 July;**13**(5):50.

Freemantle 1997

Freemantle N, Harvey EL, Wolf F, Grimshaw JM, Grilli R, Bero LA. Printed educational materials: effects on professional practice and health care outcomes (Cochrane Review). *Cochrane Database of Systematic Reviews* 2003, Issue 2. [DOI: 10.1002/14651858.CD000172]

Gardner 2010

Gardner B, Whittington C, McAteer J, Eccles MP, Michie S. Using theory to synthesise evidence from behaviour change interventions: the example of audit and feedback. *Social Science in Medicine* 2010 May;**70**(10):1618-25.

Grimshaw 2001

Grimshaw JM, Shirran L, Thomas R, Mowatt G, Fraser C, Bero L. Changing provider behavior: An overview of systematic reviews of interventions. *Medical Care* 2001;**39**(Suppl 2):II-2 - II-45.

Grimshaw 2004

Grimshaw JM, Thomas RE, MacLennan G, Fraser C, Ramsay CR, Vale L, et al. Effectiveness and efficiency of guideline dissemination and implementation strategies. *Health Technology Assessment* 2004;**8**(6):iii-iv, 1-72.

Grol 2007

Grol RP, Bosch MC, Hulscher ME, Eccles MP, Wensing M. Planning and studying improvement in patient care: the use of theoretical perspectives. *Milbank Quarterly* 2007;**85**(1):93-138.

Guldberg 2009

Guldberg TL, Lauritzen T, Kristensen JK, Vedsted P. The effect of feedback to general practitioners on quality of care for people with type 2 diabetes. A systematic review of the literature. *BMC Family Practice* 2009;**10**:30.

Guyatt 2008

Guyatt GH, Oxman AD, Kunz R, Vist GE, Falck-Ytter Y, Schunemann HJ, and the GRADE Working Group. What is 'quality of evidence' and why is it important to clinicians?. *BMJ* 2008;**336**:995-8.

Higgins 2008

Higgins JPT, Green S (editors). Cochrane Handbook for Systematic Reviews of Interventions. Chichester (UK): Wiley-Blackwell, 2008.

Hysong 2006

Hysong SJ, Best RG, Pugh JA. Audit and feedback and clinical practice guideline adherence: Making feedback actionable. *Implementation Science* 2006 April;**28**(1):9.

Hysong 2009

Hysong SJ. Meta-Analysis: audit and feedback features impact effectiveness on care quality. *Medical Care* 2009 March;**47**(3):356-63.

Kluger 1996

Kluger AN, DeNisi A. The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin* 1996;**119**:254-84.

Locke 2002

Locke EA, Latham GP. Building a practically useful theory of goal setting and task motivation. A 35-year odyssey. *American Psychologist* 2002;**57**(9):705-17.

Mugford 1991

Mugford M, Banfield P, O'Hanlon M. Effects of feedback of information on clinical practice: a review. *BMJ* 1991;**303**:398-402.

Nasser 2008

Nasser M, Oxman AD, Paulsen E, Fedorowicz Z. Local consensus processes: effects on professional practice and health care outcomes. *Cochrane Database of Systematic Reviews* 2007, Issue 1. [DOI: 10.1002/14651858.CD003165.pub3]

O'Brien 2008

O'Brien MA, Rogers S, Jamtvedt G, Oxman AD, Odgaard-Jensen J, Kristoffersen DT, et al. Educational outreach visits: effects on professional practice and health care outcomes. *Cochrane Database of Systematic Reviews* 2007, Issue 4. [DOI: 10.1002/14651858.CD000409.pub2]

Odgard-Jensen 2011

Odgaard-Jensen J, Vist GE, Timmer A, Kunz R, Akl EA, Schunemann H, et al. Randomisation to protect against selection bias in healthcare trials. *Cochrane Database of Systematic Reviews* 2011, Issue 4. [DOI: 10.1002/14651858.MR000012.pub2]

Schunemann 2008

Schünemann HJ, Oxman AD, Vist GE, Higgins JPT, Deeks JJ, Glasziou P, et al. Interpreting results and drawing conclusions. Chapter 12. In: Higgins JPT, Green S editor(s). Cochrane Handbook for Systematic Reviews of Interventions. Chichester (UK): Wiley-Blackwell, 2008.

Schunemann 2009

Schünemann H, Brożek J, Oxman A, editors. GRADE handbook for grading quality of evidence and strength of recommendation. *The GRADE Working Group.* 2009 Version 3.2 [updated March 2009];**Available from**:http://www.cc-ims.net/gradepro.

Shojania 2006

Shojania KG, Ranji SR, McDonald KM, Grimshaw JM, Sundaram V, Rushakoff RJ, et al. Effects of quality improvement strategies for type 2 diabetes on glycemic control: a meta-regression analysis.. *JAMA* 2006;**296**(4):427-40.

Shute 2008

Shute V. Focus on Formative Feedback. *Review of Educational Research* 2008;**78**(153):DOI: 10.3102/0034654307313795.

Simera 2010

Simera I, Moher D, Hirst A, Hoey J, Schulz KF, Altman DG. Transparent and accurate reporting increases reliability, utility, and impact of your research: reporting guidelines and the EQUATOR Network. *BMC Medicine* 2010;**8**:24.

Sniehotta 2009

Sniehotta FF. Towards a theory of intentional behaviour change: Plans, planning, and self-regulation. *British Journal of Health Psychology* 2009;**14**:261-73.

Van der Veer 2010

Van der Veer SN, De Keizer NF, Ravelli AC, Tenkink S, Jager KJ. Improving quality of care. A systematic review on how medical registries provide information feedback to health care providers. *International Journal of Medical Informatics* 2010 May;**79**(5):305-23.

References to other published versions of this review

Jamtvedt 2003

Jamtvedt G, Young JM, Kristoffersen DT, OBrian MA, Oxman A. Audit and feedback: effects on professional practice and health care outcomes. *Cochrane Database of Systematic Reviews* 2006, Issue 2. [DOI: 10.1002/14651858.CD000259.pub2]

Jamtvedt 2006

Jamtvedt G, Young JM, Kristoffersen DT, O'Brien MA, Oxman AD. Audit and feedback: effects on professional practice and health care outcomes.. *Cochrane Database of Systematic Reviews* 2006, Issue 2. [DOI: 10.1002/14651858.CD000259.pub2]

Thomson OBrien 1997a

Thomson O'Brien MA, Oxman AD, Davis DA, Haynes RB, Freemantle N, Harvey EL. Audit and feedback versus alternative strategies: effects on professional practice and health care outcomes. *Cochrane Database of Systematic Reviews* 2003, Issue 1. [DOI: 10.1002/14651858.CD000260]

Thomson OBrien 1997b

Thomson O'Brien MA, Oxman AD, Davis DA, Haynes RB, Freemantle N, Harvey EL. Audit and feedback: effects on professional practice and health care outcomes. *Cochrane Database of Systematic Reviews* 2003, Issue 1. [DOI: 10.1002/14651858.CD000259.pub2]

* Indicates the major publication for the study

CHARACTERISTICS OF STUDIES

Characteristics of included studies *[ordered by study ID]*

Anderson 1994

Methods	Design: cluster RCT
Participants	Country: USA Setting: inpatient Specialty: internists N health professionals: 646 N patients: -
Interventions	Description of Groups: usual care vs AF(group) + education vs AF(group + ind) + education
Outcomes	Targeted behaviour: compliance with guidelines for DVT Baseline performance: unclear
Notes	Format: verbal and written Source: unclear Frequency: unclear Instructions: no explicit target or action plan

Anderson 1994 *(Continued)*

	Nature of change: increase

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Described as "drawing lots"
Allocation concealment (selection bias)	Low risk	Cluster trial, allocation after recruitment completed
Blinding (performance bias and detection bias) All outcomes	Unclear risk	Unable to assess
Incomplete outcome data (attrition bias) All outcomes	Low risk	All hospitals followed up
Selective reporting (reporting bias)	Low risk	Appropriate outcomes measured
Baseline similar?	Unclear risk	Unable to assess
No contamination?	Low risk	No evidence of contamination
Risk of bias overall?	Unclear risk	Unable to assess

Avery 2010

Methods	Design: cluster RCT
Participants	Country: UK Setting: outpatient Specialty: gp / family physician N health professionals: unclear (72 practices) N patients: unclear
Interventions	Description of Groups: feedback vs feedback with educational outreach by pharmacists
Outcomes	Targeted behaviour: safe prescribing of NSAIDs, ACE, BB (3) Baseline performance: high
Notes	Format: written Source: investigators Frequency: once only Instructions: no explicit target, but action plan given Nature of change: increase

Avery 2010 *(Continued)*

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Block randomization with stratification
Allocation concealment (selection bias)	Low risk	Cluster trial, allocation after recruitment completed
Blinding (performance bias and detection bias) All outcomes	Low risk	Data collected from database
Incomplete outcome data (attrition bias) All outcomes	Low risk	No practices lost to follow-up
Selective reporting (reporting bias)	Low risk	Appropriate outcomes measured
Baseline similar?	Low risk	Groups similar, see Table 3
No contamination?	Low risk	No evidence of contamination
Risk of bias overall?	Low risk	

Awad 2006

Methods	Design: cluster RCT
Participants	Country: Sudan Setting: primary care or outpatient Specialty: GP/family physicians N health professionals: 50 N patients: 1800
Interventions	Description of Groups: usual care vs AF w reminders vs AF w reminders and seminar vs AF w reminders and education via academic detailing
Outcomes	Targeted behaviour: antibiotic rx Baseline performance: moderate
Notes	Format: verbal and written Source: investigators Frequency: monthly Instructions: action plan provided, but no specific target Nature of change: decrease

Awad 2006 *(Continued)*

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Unable to assess
Allocation concealment (selection bias)	Unclear risk	Unable to assess
Blinding (performance bias and detection bias) All outcomes	Unclear risk	Unable to assess
Incomplete outcome data (attrition bias) All outcomes	Low risk	Data collected from all 20 health centres
Selective reporting (reporting bias)	Unclear risk	Unable to assess
Baseline similar?	Low risk	Baseline encounters and prescriptions similar
No contamination?	Low risk	No evidence of contamination
Risk of bias overall?	Unclear risk	Unable to assess

Bahrami 2004

Methods	Design: cluster RCT
Participants	Country: Scotland Setting: primary care or outpatient Specialty: dentists N health professionals: 51 N patients: 1934
Interventions	Description of Groups: guideline vs guideline plus AF vs guideline plus computer decision support vs all
Outcomes	Targeted behaviour: compliance w guideline for impacted molars Baseline performance: high
Notes	Format: unclear Source: investigators Frequency: once only Instructions: no explicit target or action plan Nature of change: decrease

Bahrami 2004 *(Continued)*

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Computer random number table
Allocation concealment (selection bias)	Low risk	Cluster trial, allocation after recruitment completed
Blinding (performance bias and detection bias) All outcomes	Low risk	Data extractor was blinded
Incomplete outcome data (attrition bias) All outcomes	Low risk	Less than 10% drop out, spread among groups
Selective reporting (reporting bias)	Unclear risk	Unable to assess
Baseline similar?	Low risk	See Table 1
No contamination?	Low risk	Separate practices
Risk of bias overall?	Low risk	

Baker 1997

Methods	Design: cluster RCT
Participants	Country: UK Setting: primary care or outpatient Specialty: GP/family physicians N health professionals: 18 practices N patients: 2409
Interventions	Description of Groups: AF vs AF + reminders
Outcomes	Targeted behaviour: appropriate benzodiazepine use Baseline performance: low
Notes	Format: written Source: unclear Frequency: once only Instructions: no explicit target or action plan Nature of change: increase

Baker 1997 *(Continued)*

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Random number table
Allocation concealment (selection bias)	Low risk	Cluster trial, allocation after recruitment completed
Blinding (performance bias and detection bias) All outcomes	High risk	Unable to blind second abstractor
Incomplete outcome data (attrition bias) All outcomes	Low risk	Data from all 18 practices
Selective reporting (reporting bias)	Low risk	Appropriate outcomes reported
Baseline similar?	Low risk	See Table 1
No contamination?	Low risk	Separate practices
Risk of bias overall?	High risk	Non-blinded outcome

Baker 2003

Methods	Design: cluster RCT
Participants	Country: UK Setting: primary care or outpatient Specialty: GP/family physicians N health professionals: 96 N patients: -
Interventions	Description of Groups: AF vs control
Outcomes	Targeted behaviour: lipid screening Baseline performance: unclear
Notes	Format: written Source: investigators Frequency: less than monthly, more than once Instructions: no explicit target or action plan Nature of change: decrease

Baker 2003 *(Continued)*

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Random number table
Allocation concealment (selection bias)	Low risk	Cluster trial, allocation after recruitment completed
Blinding (performance bias and detection bias) All outcomes	Unclear risk	Unable to assess
Incomplete outcome data (attrition bias) All outcomes	Low risk	No sites lost to follow-up
Selective reporting (reporting bias)	Low risk	Appropriate outcomes reported
Baseline similar?	Unclear risk	Unable to assess
No contamination?	Low risk	GPs work separately
Risk of bias overall?	Unclear risk	Unable to assess

Baker 2003A

Methods	Design: cluster RCT
Participants	Country: UK Setting: primary care or outpatient Specialty: GP/family physicians N health professionals: 225 N patients: -
Interventions	Description of Groups: usual care vs education vs AF
Outcomes	Targeted behaviour: compliance with guidelines for asthma and angina Baseline performance: unclear
Notes	Format: written Source: investigators Frequency: once only Instructions: no explicit target or action plan Nature of change: increase

Baker 2003A *(Continued)*

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Computer generated for allocation
Allocation concealment (selection bias)	Low risk	Cluster trial, allocation after recruitment completed
Blinding (performance bias and detection bias) All outcomes	Low risk	Data collectors blinded
Incomplete outcome data (attrition bias) All outcomes	Low risk	All practices completed study
Selective reporting (reporting bias)	Low risk	Appropriate outcomes reported
Baseline similar?	Low risk	See Table 2
No contamination?	Low risk	Separate practices
Risk of bias overall?	Low risk	

Balas 1998

Methods	Design: cluster RCT
Participants	Country: USA Setting: primary care or outpatient Specialty: internists N health professionals: 10 N patients: 152
Interventions	Description of Groups: AF vs control
Outcomes	Targeted behaviour: use of peritoneal dialysis (rather than hemodialysis) Baseline performance: unclear
Notes	Format: written Source: unclear Frequency: monthly Instructions: no explicit target or action plan Nature of change: mix of increase and decrease or change behaviours

Balas 1998 *(Continued)*

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Computer generated randomization
Allocation concealment (selection bias)	Low risk	Cluster trial, allocation after recruitment completed
Blinding (performance bias and detection bias) All outcomes	Low risk	Data collectors blinded
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Unable to assess
Selective reporting (reporting bias)	Low risk	Appropriate outcomes reported
Baseline similar?	Unclear risk	Unable to assess
No contamination?	Unclear risk	Unable to assess
Risk of bias overall?	Unclear risk	Unable to assess

Batty 2001

Methods	Design: cluster RCT
Participants	Country: UK Setting: inpatient Specialty: internists N health professionals: 70 (17 hospitals) N patients: 539
Interventions	Description of Groups: usual care vs AF verbal vs AF written
Outcomes	Targeted behaviour: appropriate prescribing of benzodiazepines Baseline performance: moderate
Notes	Format: both verbal and written Source: investigators Frequency: once only Instructions: no explicit target or action plan Nature of change: decrease

Batty 2001 *(Continued)*

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Unable to assess
Allocation concealment (selection bias)	Low risk	Cluster trial, allocation after recruitment completed
Blinding (performance bias and detection bias) All outcomes	Unclear risk	Unable to assess
Incomplete outcome data (attrition bias) All outcomes	Low risk	Data from all 17 hospitals
Selective reporting (reporting bias)	Low risk	Appropriate outcomes reported
Baseline similar?	High risk	Figure 2, variaility at baseline between groups
No contamination?	Unclear risk	Unable to assess
Risk of bias overall?	High risk	Baseline variability

Beck 2005

Methods	Design: Cluster RCT
Participants	Country: Canada Setting: inpatient Specialty: internists N health professionals: unclear (76 hospitals) N patients: 5675
Interventions	Description of Groups: AF vs control
Outcomes	Targeted behaviour: prescribing beta-blockers Baseline performance: moderate
Notes	Format: written Source: investigators Frequency: once only Instructions: specific target, but not action plan Nature of change: increase

Beck 2005 *(Continued)*

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Computer generated randomization
Allocation concealment (selection bias)	Low risk	Cluster trial, allocation after recruitment completed
Blinding (performance bias and detection bias) All outcomes	Low risk	Data collected from database
Incomplete outcome data (attrition bias) All outcomes	Low risk	No evidence of attrition
Selective reporting (reporting bias)	Unclear risk	Unable to assess
Baseline similar?	Unclear risk	Unable to assess
No contamination?	Low risk	No evidence of contamination
Risk of bias overall?	Low risk	

Bentz 2007

Methods	Design: cluster RCT
Participants	Country: USA Setting: primary care or outpatient Specialty: GP/family physicians (and nurses) N health professionals: 279 N patients: 102915
Interventions	Description of Groups: education vs clinical decision support plus feedback
Outcomes	Targeted behaviour: smoking cessation referrals Baseline performance: moderate
Notes	Format: verbal and written Source: investigators Frequency: monthly Instructions: no explicit target or action plan Nature of change: increase

Bentz 2007 *(Continued)*

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Quote "clinics matched and randomized"
Allocation concealment (selection bias)	Low risk	Cluster trial, recruitment not influenced by allocation
Blinding (performance bias and detection bias) All outcomes	Unclear risk	Unable to assess
Incomplete outcome data (attrition bias) All outcomes	Low risk	No evidence of attrition
Selective reporting (reporting bias)	Low risk	Appropriate outcomes accounted for
Baseline similar?	Low risk	Clinics comparable, see Table 1
No contamination?	Low risk	No evidence of contamination
Risk of bias overall?	Low risk	

Berman 1998

Methods	Design: patient or provider level RCT
Participants	Country: USA Setting: inpatient Specialty: anasthesiologists N health professionals: 27 N patients: -
Interventions	Description of Groups: AF vs control
Outcomes	Targeted behaviour: use of high cost anesthetic drugs Baseline performance: unclear
Notes	Format: written Source: unclear Frequency: more than monthly Instructions: no explicit target or action plan Nature of change: decrease

Berman 1998 *(Continued)*

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Quote "randomized into two groups"
Allocation concealment (selection bias)	Low risk	Cluster trial, allocation after recruitment completed
Blinding (performance bias and detection bias) All outcomes	Unclear risk	Unable to assess
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Unable to assess
Selective reporting (reporting bias)	Unclear risk	Unable to assess
Baseline similar?	High risk	Differences between groups, see Table 1
No contamination?	High risk	Unable to rule out contamination as groups work closely together
Risk of bias overall?	High risk	Baseline differences

Blais 2008

Methods	Design: patient or provider level RCT
Participants	Country: Canada Setting: primary care or outpatient Specialty: GP/family physicians, OBGYN N health professionals: 131 N patients: -
Interventions	Description of Groups: usual care vs feedback
Outcomes	Targeted behaviour: compliance w guideline for asthma Baseline performance: high
Notes	Format: written Source: investigators Frequency: less than monthly, more than once Instructions: no explicit target or action plan Nature of change: mix of increase and decrease or change behaviours

Blais 2008 *(Continued)*

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Unable to assess
Allocation concealment (selection bias)	Low risk	Cluster trial, allocation after recruitment completed
Blinding (performance bias and detection bias) All outcomes	Low risk	Administrative data used
Incomplete outcome data (attrition bias) All outcomes	Low risk	Dropouts comparable in both groups (page 229)
Selective reporting (reporting bias)	Low risk	Appropriate outcomes reported
Baseline similar?	Low risk	Table 2 and Table 3
No contamination?	Unclear risk	Unable to assess
Risk of bias overall?	Low risk	

Boekeloo 1990

Methods	Design: Cluster RCT
Participants	Country: USA Setting: inpatient Specialty: internists N health professionals: 29 N patients: -
Interventions	Description of Groups: usual care vs reminders vs feedback vs both
Outcomes	Targeted behaviour: compliance with cholesterol guidelines Baseline performance: low
Notes	Format: written Source: supervisor Frequency: once only Instructions: action plan provided, but no specific target Nature of change: increase

Boekeloo 1990 *(Continued)*

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Unable to assess
Allocation concealment (selection bias)	Low risk	Cluster trial, allocation after recruitment completed
Blinding (performance bias and detection bias) All outcomes	Unclear risk	Unable to assess
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Unable to assess
Selective reporting (reporting bias)	Unclear risk	Unable to assess
Baseline similar?	High risk	Patients differed in key variables (lab values, medical history)
No contamination?	High risk	Contamination between physicians likely
Risk of bias overall?	High risk	Baseline data different

Bonevski 1999

Methods	Design: Cluster RCT
Participants	Country: Australia Setting: primary care or outpatient Specialty: GP/family physicians N health professionals: 19 N patients: -
Interventions	Description of Groups: guidelines vs guidelines plus feedback
Outcomes	Targeted behaviour: screening for cholesterol and bp Baseline performance: unclear
Notes	Format: written Source: investigators Frequency: once only Instructions: explicit, measurable target and action plan Nature of change: increase

Bonevski 1999 *(Continued)*

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Unable to assess
Allocation concealment (selection bias)	Unclear risk	Unable to assess
Blinding (performance bias and detection bias) All outcomes	Low risk	Patients appear to be unaware of intervention arm
Incomplete outcome data (attrition bias) All outcomes	Low risk	High participation rates, no evidence of variation between groups
Selective reporting (reporting bias)	Low risk	Appropriate outcomes reported
Baseline similar?	Unclear risk	Unable to assess
No contamination?	Unclear risk	Unable to assess
Risk of bias overall?	Unclear risk	Unable to assess

Borgiel 1999

Methods	Design: patient or provider level RCT
Participants	Country: Canada Setting: primary care or outpatient Specialty: GP/family physicians N health professionals: 56 N patients: -
Interventions	Description of Groups: AF vs AF + educational outreach
Outcomes	Targeted behaviour: quality of care scores for prevention Baseline performance: high
Notes	Format: written Source: rep from employer or quality assurance org Frequency: once only Instructions: no explicit target or action plan Nature of change: increase

Borgiel 1999 *(Continued)*

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Stratification and block randomization
Allocation concealment (selection bias)	Unclear risk	Unable to assess
Blinding (performance bias and detection bias) All outcomes	Unclear risk	Unable to assess
Incomplete outcome data (attrition bias) All outcomes	Low risk	Equal numbers of subjects in both groups completed study
Selective reporting (reporting bias)	Unclear risk	Unable to assess
Baseline similar?	Low risk	Physicians and patients comparable (Table 1 and 2)
No contamination?	Unclear risk	Unable to assess
Risk of bias overall?	Unclear risk	Unable to assess

Brady 1988

Methods	Design: Cluster RCT
Participants	Country: USA Setting: primary care or outpatient Specialty: internists N health professionals: 45 N patients: 255
Interventions	Description of Groups: AF vs AF + education + self-audit with active control
Outcomes	Targeted behaviour: compliance with guidelines for flu vacc and mammography Baseline performance: unclear
Notes	Format: verbal Source: unclear Frequency: once only Instructions: no explicit target or action plan Nature of change: increase

Brady 1988 *(Continued)*

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Random number table used for allocation
Allocation concealment (selection bias)	Low risk	Cluster trial, allocation after recruitment completed
Blinding (performance bias and detection bias) All outcomes	Unclear risk	Unable to assess
Incomplete outcome data (attrition bias) All outcomes	Low risk	All residents were followed up
Selective reporting (reporting bias)	Low risk	Appropriate outcomes reported
Baseline similar?	Unclear risk	Unable to assess
No contamination?	Unclear risk	Unable to assess
Risk of bias overall?	Unclear risk	Unable to assess

Bregnhoj 2009

Methods	Design: Cluster RCT
Participants	Country: Denmark Setting: primary care or outpatient Specialty: GP/family physicians N health professionals: 41 N patients: 212
Interventions	Description of Groups: usual care vs educational meeting only vs educational meeting plus AF
Outcomes	Targeted behaviour: medication appropriateness index Baseline performance: unclear
Notes	Format: verbal and written Source: investigators Frequency: once only Instructions: action plan provided, but no specific target Nature of change: mix of increase and decrease or change behaviours

Bregnhoj 2009 *(Continued)*

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Computer generated randomization list
Allocation concealment (selection bias)	Low risk	Cluster trial, allocation after recruitment completed
Blinding (performance bias and detection bias) All outcomes	Low risk	Quote "Evaluators blinded"
Incomplete outcome data (attrition bias) All outcomes	Low risk	GPs followed up equally; slight differences in number of patients, but likely not due to intervention
Selective reporting (reporting bias)	Low risk	Appropriate outcomes collected
Baseline similar?	Low risk	See Table 1 and 2
No contamination?	Low risk	GPs in separate practices
Risk of bias overall?	Low risk	

Brown 1994

Methods	Design: cluster RCT
Participants	Country: Australia Setting: primary care or outpatient Specialty: dentists N health professionals: 24 practices N patients: -
Interventions	Description of Groups: AF with educational outreach vs control
Outcomes	Targeted behaviour: recording of periodontal care Baseline performance: high
Notes	Format: verbal and written Source: unclear Frequency: once only Instructions: no explicit target or action plan Nature of change: increase

Brown 1994 *(Continued)*

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Unable to assess
Allocation concealment (selection bias)	Low risk	Cluster trial, allocation after recruitment completed
Blinding (performance bias and detection bias) All outcomes	Unclear risk	Unable to assess
Incomplete outcome data (attrition bias) All outcomes	Low risk	Minimal number of practices lost to follow-up
Selective reporting (reporting bias)	Low risk	Appropriate outcomes measured
Baseline similar?	High risk	Groups different at baseline (see Table 1)
No contamination?	Low risk	Practices separate
Risk of bias overall?	High risk	Baseline variability

Buffington 1991

Methods	Design: cluster RCT
Participants	Country: USA Setting: primary care or outpatient Specialty: GP/family physicians, internists N health professionals: 45 N patients: -
Interventions	Description of Groups: usual care vs AF vs AF + patient reminders
Outcomes	Targeted behaviour: influenza vacc rates Baseline performance: unclear
Notes	Format: written Source: unclear Frequency: more than monthly Instructions: no explicit target or action plan Nature of change: increase

Buffington 1991 *(Continued)*

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Unable to assess
Allocation concealment (selection bias)	Low risk	Cluster trial, allocation after recruitment completed
Blinding (performance bias and detection bias) All outcomes	High risk	Outcomes of physician report
Incomplete outcome data (attrition bias) All outcomes	Low risk	Quote "All practices...successfully monitored..."
Selective reporting (reporting bias)	Low risk	Appropriate outcomes measured
Baseline similar?	Unclear risk	Unable to assess
No contamination?	Unclear risk	Unable to assess
Risk of bias overall?	High risk	Physician self report of outcome

Buntinx 1993

Methods	Design: patient or provider level RCT
Participants	Country: Belgium Setting: primary care or outpatient Specialty: GP/family physicians, obgyn N health professionals: 179 N patients: -
Interventions	Description of Groups: usual care x2 vs feedback vs feedback with recommendations
Outcomes	Targeted behaviour: quality of pap smears Baseline performance: high
Notes	Format: written Source: unclear Frequency: monthly Instructions: action plan provided, but no specific target Nature of change: increase

Buntinx 1993 *(Continued)*

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Unable to assess
Allocation concealment (selection bias)	Low risk	Cluster trial, allocation after recruitment completed
Blinding (performance bias and detection bias) All outcomes	Low risk	Outcome assessors blinded
Incomplete outcome data (attrition bias) All outcomes	Low risk	Similar numbers excluded from the groups
Selective reporting (reporting bias)	Unclear risk	Unable to assess
Baseline similar?	Low risk	Groups similar at baseline
No contamination?	Unclear risk	Unable to assess
Risk of bias overall?	Unclear risk	Unable to assess

Canovas 2009

Methods	Design: cluster RCT
Participants	Country: Spain Setting: primary care or outpatient Specialty: GP/family physicians N health professionals: 10 practices N patients: -
Interventions	Description of Groups: usual care vs internal QA cycle vs feedback
Outcomes	Targeted behaviour: prescribing for common cold Baseline performance: moderate
Notes	Format: verbal Source: investigators Frequency: more than monthly Instructions: no explicit target or action plan Nature of change: mix of increase and decrease or change behaviours

Canovas 2009 *(Continued)*

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Unable to assess
Allocation concealment (selection bias)	Low risk	Cluster trial, allocation after recruitment completed
Blinding (performance bias and detection bias) All outcomes	Unclear risk	Unable to assess
Incomplete outcome data (attrition bias) All outcomes	High risk	Poor follow-up
Selective reporting (reporting bias)	Low risk	Appropriate outcomes measured
Baseline similar?	High risk	Table 2
No contamination?	Low risk	Randomized by centre, unlikely to be contamination
Risk of bias overall?	High risk	Differences at baseline, dropouts in intervention, unclear randomization

Charrier 2008

Methods	Design: cluster RCT
Participants	Country: Italy Setting: inpatient Specialty: nurses in mixed depts in hospitals N health professionals: 160 N patients: -
Interventions	Description of Groups: usual care vs audit and feedback with facilitators
Outcomes	Targeted behaviour: compliance w protocols for venous catheters and pressure ulcers Baseline performance: moderate
Notes	Format: unclear Source: investigators Frequency: less than monthly, more than once Instructions: no explicit target or action plan Nature of change: mix of increase and decrease or change behaviours

Charrier 2008 *(Continued)*

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Unable to assess
Allocation concealment (selection bias)	Low risk	Cluster trial, allocation after recruitment completed
Blinding (performance bias and detection bias) All outcomes	High risk	No blinding, subjective measurement
Incomplete outcome data (attrition bias) All outcomes	Low risk	All units followed up
Selective reporting (reporting bias)	Low risk	Appropriate outcomes measured
Baseline similar?	Unclear risk	Unable to assess
No contamination?	Low risk	No evidence of contamination, units separate
Risk of bias overall?	High risk	Variable balance at baseline and multiple testing and evaluation not blinded

Chassin 1986

Methods	Design: cluster RCT
Participants	Country: USA Setting: inpatient Specialty: obgyn N health professionals: 1483 N patients: -
Interventions	Description of Groups: feedback + education vs control
Outcomes	Targeted behaviour: antenatal xray pelvimetry rates Baseline performance: unclear
Notes	Format: written Source: supervisor Frequency: monthly Instructions: no explicit target or action plan Nature of change: decrease

Chassin 1986 *(Continued)*

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Unable to assess
Allocation concealment (selection bias)	Low risk	Cluster trial, allocation after recruitment completed
Blinding (performance bias and detection bias) All outcomes	Unclear risk	Unable to assess
Incomplete outcome data (attrition bias) All outcomes	Low risk	No evidence of attrition
Selective reporting (reporting bias)	Low risk	Approprate outcomes measured
Baseline similar?	Unclear risk	Unable to assess
No contamination?	Low risk	Hospitals separate, no evidence of contamination
Risk of bias overall?	Unclear risk	Unable to assess

Cheater 2006

Methods	Design: cluster RCT
Participants	Country: England Setting: primary care or outpatient Specialty: nurses in primary care N health professionals: 176 N patients: 1078
Interventions	Description of Groups: usual care v AF alone vs educational outreach vs both
Outcomes	Targeted behaviour: compliance w guideline for urinary incontinence Baseline performance: moderate
Notes	Format: written Source: investigators Frequency: once only Instructions: action plan provided, but no specific target Nature of change: mix of increase and decrease or change behaviours

Cheater 2006 *(Continued)*

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Stratified randomization, blocks of 4
Allocation concealment (selection bias)	Low risk	Quote "concealed randomization"
Blinding (performance bias and detection bias) All outcomes	Low risk	Quote "data collectors blind to allocation"
Incomplete outcome data (attrition bias) All outcomes	Low risk	All practices followed up
Selective reporting (reporting bias)	Low risk	Appropriate outcomes measured
Baseline similar?	Low risk	Groups comparable (page 545)
No contamination?	Low risk	No evidence of contamination
Risk of bias overall?	Low risk	

Claes 2005

Methods	Design: cluster RCT
Participants	Country: Belgium Setting: primary care or outpatient Specialty: family medicine / GP N health professionals: 96 (66 practices) N patients: 834
Interventions	Description of groups: usual care (with education) vs feedback vs facilitated relay with point of care testing vs computer-decision support
Outcomes	Targeted behaviour: proportion of time within target for INR Baseline performance: moderate
Notes	Format: unclear whether written or verbal Source: investigators Frequency: less than monthly Instructions: no target or action plan Nature of change: mix of increase and decrease or change behaviours

Claes 2005 *(Continued)*

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	High risk	Unusual procedure similar to drawing lots, but sequence not clearly random
Allocation concealment (selection bias)	Unclear risk	Unable to assess
Blinding (performance bias and detection bias) All outcomes	Low risk	Data collected separately
Incomplete outcome data (attrition bias) All outcomes	Low risk	Patients who did not complete study were removed for reasons not related to intervention
Selective reporting (reporting bias)	Low risk	Appropriate outcomes measured
Baseline similar?	Low risk	See Figure 2
No contamination?	Low risk	Separate practices
Risk of bias overall?	High risk	Sequence seems not random

Cline 2007

Methods	Design: cluster RCT
Participants	Country: USA Setting: primary care or outpatient Specialty: ER N health professionals: 30 N patients: -
Interventions	Description of Groups: usual care vs AF
Outcomes	Targeted behaviour: hypertension referrals Baseline performance: low
Notes	Format: written Source: investigators Frequency: more than monthly Instructions: no explicit target or action plan Nature of change: increase

Cline 2007 *(Continued)*

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Random number table
Allocation concealment (selection bias)	Low risk	Cluster trial, allocation after recruitment completed
Blinding (performance bias and detection bias) All outcomes	Unclear risk	Unable to assess
Incomplete outcome data (attrition bias) All outcomes	Low risk	Appears to have full follow up
Selective reporting (reporting bias)	Low risk	Appropriate outcomes reported
Baseline similar?	Low risk	Patient referral rates comparable
No contamination?	Unclear risk	Unable to assess
Risk of bias overall?	Low risk	

Cohen 1982

Methods	Design: cluster RCT
Participants	Country: USA Setting: inpatient Specialty: internists N health professionals: 4 practices N patients: 511
Interventions	Description of Groups: feedback vs control with active control
Outcomes	Targeted behaviour: test ordering Baseline performance: unclear
Notes	Format: written Source: unclear Frequency: more than monthly Instructions: no explicit target or action plan Nature of change: decrease

Cohen 1982 *(Continued)*

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Unable to assess
Allocation concealment (selection bias)	Low risk	Cluster trial, allocation after recruitment completed
Blinding (performance bias and detection bias) All outcomes	Low risk	Lab test collected separately
Incomplete outcome data (attrition bias) All outcomes	Low risk	Groups remained equal in size throughout study
Selective reporting (reporting bias)	Low risk	Appropriate outcomes reported
Baseline similar?	Low risk	Groups comparable at baseline
No contamination?	Low risk	Separate clinics
Risk of bias overall?	Unclear risk	Unable to assess

Curran 2008

Methods	Design: cluster RCT
Participants	Country: Scotland Setting: inpatient Specialty: nurses in mixed depts in hospitals N health professionals: 24 hospitals N patients: -
Interventions	Description of Groups: control group (but did monthly audits), feedback, and feedback with pareto charts (correct solution info)
Outcomes	Targeted behaviour: mrsa rates Baseline performance: high
Notes	Format: written Source: investigators Frequency: monthly Instructions: action plan provided, but no specific target Nature of change: mix of increase and decrease or change behaviours

Curran 2008 *(Continued)*

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Random number
Allocation concealment (selection bias)	Low risk	Cluster trial, allocation after recruitment completed
Blinding (performance bias and detection bias) All outcomes	Unclear risk	Unable to assess
Incomplete outcome data (attrition bias) All outcomes	Low risk	No evidence of attrition
Selective reporting (reporting bias)	Low risk	Approipriate outcomes measured
Baseline similar?	Low risk	Groups comparable at baseline
No contamination?	High risk	Each hospital had one of each group
Risk of bias overall?	High risk	Contamination likely, control group DID audit, outcome assessors unclear

Curtis 2005

Methods	Design: patient or provider level RCT
Participants	Country: USA Setting: primary care or outpatient Specialty: GP/family physicians, internists N health professionals: 101 N patients: 421
Interventions	Description of Groups: usual care vs feedback with education
Outcomes	Targeted behaviour: improvement in safe nsaid prescribing Baseline performance: moderate
Notes	Format: written Source: investigators Frequency: less than monthly, more than once Instructions: explicit, measurable target but no action plan Nature of change: increase

Curtis 2005 *(Continued)*

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Unable to assess
Allocation concealment (selection bias)	Low risk	Cluster trial, allocation after recruitment completed
Blinding (performance bias and detection bias) All outcomes	Low risk	Data abstractors blind to treatment
Incomplete outcome data (attrition bias) All outcomes	Low risk	No difference between groups
Selective reporting (reporting bias)	Low risk	Appropriate outcomes measured
Baseline similar?	Low risk	See Table 1
No contamination?	Unclear risk	Unable to assess
Risk of bias overall?	Unclear risk	Unable to assess

Curtis 2007

Methods	Design: patient or provider level RCT
Participants	Country: USA Setting: primary care or outpatient Specialty: GP/family physicians, internists N health professionals: 153 N patients: 949
Interventions	Description of Groups: usual care vs web-based education modules and feedback
Outcomes	Targeted behaviour: osteoporosis management Baseline performance: moderate
Notes	Format: unclear Source: investigators Frequency: once only Instructions: explicit, measurable target but no action plan Nature of change: increase

Curtis 2007 *(Continued)*

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Unable to assess
Allocation concealment (selection bias)	Unclear risk	Unable to assess
Blinding (performance bias and detection bias) All outcomes	Low risk	Administrative data
Incomplete outcome data (attrition bias) All outcomes	Low risk	Poor follow up in both groups, but not statistically different
Selective reporting (reporting bias)	Low risk	Appropriate outcomes measured
Baseline similar?	Low risk	Groups comparable, see page 593
No contamination?	Low risk	No evidence of contamination
Risk of bias overall?	Low risk	

De Almeida Neto 2000

Methods	Design: cluster RCT
Participants	Country: Australia Setting: primary care or outpatient Specialty: pharmacists N health professionals: 24 N patients: -
Interventions	Description of Groups: feedback + education vs control
Outcomes	Targeted behaviour: analgesic misuse identified and discussed Baseline performance: unclear
Notes	Format: verbal Source: investigators Frequency: unclear Instructions: action plan provided, but no specific target Nature of change: mix of increase and decrease or change behaviours

De Almeida Neto 2000 *(Continued)*

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Unable to assess
Allocation concealment (selection bias)	Low risk	Cluster trial, allocation after recruitment completed
Blinding (performance bias and detection bias) All outcomes	Unclear risk	Unable to assess
Incomplete outcome data (attrition bias) All outcomes	Low risk	2 pharmacists dropped out of study
Selective reporting (reporting bias)	Low risk	Appropriate outcomes measured
Baseline similar?	Low risk	See Table 1
No contamination?	Low risk	Community pharmacists
Risk of bias overall?	Unclear risk	Unable to assess

Eccles 2001

Methods	Design: cluster RCT
Participants	Country: UK Setting: primary care or outpatient Specialty: GP/family physicians N health professionals: 162 N patients: 788
Interventions	Description of Groups: usual care vs feedback vs reminders vs both
Outcomes	Targeted behaviour: number of radiograph requested for knee and lunbal spine/compliance with guideline Baseline performance: unclear
Notes	Format: written Source: investigators Frequency: less than monthly, more than once Instructions: no explicit target or action plan Nature of change: decrease

Eccles 2001 *(Continued)*

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Random number table
Allocation concealment (selection bias)	Low risk	Cluster trial, allocation after recruitment completed
Blinding (performance bias and detection bias) All outcomes	Low risk	Radiology departments not aware of randomizations
Incomplete outcome data (attrition bias) All outcomes	Low risk	Minimal dropout
Selective reporting (reporting bias)	Low risk	Appropriate outcomes measured
Baseline similar?	Unclear risk	Unable to assess
No contamination?	Low risk	Randomized by practice
Risk of bias overall?	Unclear risk	Unable to assess

Eltayeb 2005

Methods	Design: cluster RCT
Participants	Country: Sudan Setting: primary care or outpatient Specialty: medical officers, medical assistants N health professionals: 37 (20 centers) N patients: 600
Interventions	Description of Groups: usual care vs AF vs AF plus educational seminar vs feedback plus educational outreach
Outcomes	Targeted behaviour: antibiotic prescriptions Baseline performance: high
Notes	Format: written Source: investigators Frequency: once only Instructions: no explicit target or action plan Nature of change: mix / unclear

Eltayeb 2005 *(Continued)*

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Unable to assess
Allocation concealment (selection bias)	Unclear risk	Unable to assess
Blinding (performance bias and detection bias) All outcomes	Unclear risk	Unable to assess
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Unable to assess
Selective reporting (reporting bias)	Unclear risk	Unable to assess
Baseline similar?	Low risk	See Table 2
No contamination?	Low risk	Health centres separate
Risk of bias overall?	Unclear risk	Unable to assess

Everett 1983

Methods	Design: cluster RCT
Participants	Country: USA Setting: inpatient Specialty: internists N health professionals: 24 N patients: 1140
Interventions	Description of Groups: feedback + education vs control
Outcomes	Targeted behaviour: use of lab tests Baseline performance: unclear
Notes	Format: verbal and written Source: supervisor Frequency: monthly Instructions: action plan provided, but no specific target Nature of change: decrease

Everett 1983 *(Continued)*

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Unable to assess
Allocation concealment (selection bias)	Unclear risk	Unable to assess
Blinding (performance bias and detection bias) All outcomes	Unclear risk	Unable to assess
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Unable to assess
Selective reporting (reporting bias)	Unclear risk	Unable to assess
Baseline similar?	High risk	See Table 1
No contamination?	Unclear risk	Unable to assess
Risk of bias overall?	High risk	Baseline

Fairbrother 1999

Methods	Design: cluster RCT
Participants	Country: USA Setting: primary care or outpatient Specialty: pediatricians, GP/family physicians N health professionals: 61 N patients: -
Interventions	Description of Groups: usual care vs feedback vs feedback + incentive for targets vs feedback + incentive per service
Outcomes	Targeted behaviour: immunisation rates Baseline performance: unclear
Notes	Format: written Source: unclear Frequency: less than monthly, more than once Instructions: explicit, measurable target but no action plan Nature of change: increase

Fairbrother 1999 *(Continued)*

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Unable to assess
Allocation concealment (selection bias)	Unclear risk	Unable to assess
Blinding (performance bias and detection bias) All outcomes	Unclear risk	Unable to assess
Incomplete outcome data (attrition bias) All outcomes	Low risk	Only one dropout
Selective reporting (reporting bias)	Low risk	Appropriate outcomes measured
Baseline similar?	High risk	Rates of immunization differ
No contamination?	Low risk	Separate practices
Risk of bias overall?	Unclear risk	Unable to assess

Ferguson 2003

Methods	Design: cluster RCT
Participants	Country: USA Setting: inpatient Specialty: surgeons (and entire hospital) N health professionals: 359 hospitals N patients: 267977
Interventions	Description of Groups: usual care x2 vs feedback + opinion leader + education with active control
Outcomes	Targeted behaviour: compliance with guidelines for use of beta-blockers, use of IMI for CABG Baseline performance: unclear
Notes	Format: written Source: investigators Frequency: less than monthly, more than once Instructions: action plan provided, but no specific target Nature of change: increase

Ferguson 2003 *(Continued)*

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Unable to assess
Allocation concealment (selection bias)	Low risk	Cluster trial, allocation after recruitment completed
Blinding (performance bias and detection bias) All outcomes	Low risk	Participants unaware
Incomplete outcome data (attrition bias) All outcomes	Low risk	Roughly equal dropouts in 3 groups
Selective reporting (reporting bias)	Low risk	Appropriate outcomes measured
Baseline similar?	Low risk	Demographics similar in 3 groups
No contamination?	Low risk	Separate hospitals
Risk of bias overall?	Low risk	

Filardo 2009

Methods	Design: cluster RCT
Participants	Country: USA Setting: inpatient Specialty: internists N health professionals: 45 hospitals N patients: -
Interventions	Description of Groups: AF vs AF plus education and continuous quality improvement
Outcomes	Targeted behaviour: composite quality indicators for heart failure and pneumonia Baseline performance: high
Notes	Format: written Source: investigators Frequency: unclear Instructions: neither action plan or explicit target Nature of change: mix / unclear

Filardo 2009 *(Continued)*

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Unable to assess
Allocation concealment (selection bias)	Low risk	Cluster trial, allocation after recruitment completed
Blinding (performance bias and detection bias) All outcomes	Unclear risk	Unable to assess
Incomplete outcome data (attrition bias) All outcomes	Low risk	2 of 47 hospitals dropped out
Selective reporting (reporting bias)	Low risk	Appropriate outcomes measured
Baseline similar?	Unclear risk	Unable to assess
No contamination?	Low risk	Separate hospitals
Risk of bias overall?	Unclear risk	Unable to assess

Foster 2007

Methods	Design: cluster RCT
Participants	Country: Scotland Setting: primary care or outpatient Specialty: GP/family physicians N health professionals: 16 practices N patients: -
Interventions	Description of Groups: usual care (delayed intervention) vs feedback with group academic detailing (incl practice action plans)
Outcomes	Targeted behaviour: asthmatics with peak flow test Baseline performance: low
Notes	Format: written Source: investigators Frequency: once only Instructions: action plan provided, but no specific target Nature of change: increase

Foster 2007 *(Continued)*

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Random number allocation
Allocation concealment (selection bias)	Low risk	Central randomization
Blinding (performance bias and detection bias) All outcomes	Unclear risk	Unable to assess
Incomplete outcome data (attrition bias) All outcomes	High risk	Lost half of intervention group
Selective reporting (reporting bias)	Low risk	Appropriate outcomes measured
Baseline similar?	Low risk	See Table 1
No contamination?	Low risk	Practices separate
Risk of bias overall?	High risk	Half of intervention group lost to follow-up

Foy 2004

Methods	Design: cluster RCT
Participants	Country: Scotland Setting: inpatient Specialty: obgyn N health professionals: 26 hospitals N patients: -
Interventions	Description of Groups: usual care vs feedback with education
Outcomes	Targeted behaviour: compliance w guideline for induced abortion Baseline performance: unclear
Notes	Format: verbal and written Source: rep from employer or quality assurance org Frequency: once only Instructions: action plan provided, but no specific target Nature of change: mix of increase and decrease or change behaviours

Foy 2004 *(Continued)*

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Randomized pairs by independent statistitian
Allocation concealment (selection bias)	Low risk	Cluster trial, allocation after recruitment completed
Blinding (performance bias and detection bias) All outcomes	Unclear risk	Unable to assess
Incomplete outcome data (attrition bias) All outcomes	Low risk	1 unit with no cases, otherwise complete data
Selective reporting (reporting bias)	Low risk	Appropriate outcomes measured
Baseline similar?	Low risk	See Table 1
No contamination?	Low risk	Separate hospitals
Risk of bias overall?	Low risk	

Frijiling 2002

Methods	Design: cluster RCT
Participants	Country: Netherlands Setting: primary care or outpatient Specialty: GP/family physician N health professionals: 185 N patients: -
Interventions	Description of Groups: feedback + outreach vs control
Outcomes	Targeted behaviour: % compliance with diabetes guidelines Baseline performance: unclear
Notes	Format: verbal and written Source: investigators Frequency: more than monthly Instructions: action plan provided, but no specific target Nature of change: mix of increase and decrease or change behaviours

Frijiling 2002 *(Continued)*

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Random number generator, blocks of 4
Allocation concealment (selection bias)	Low risk	Quote "person responsible for the randomization process was blind to identities"
Blinding (performance bias and detection bias) All outcomes	High risk	Physician report data
Incomplete outcome data (attrition bias) All outcomes	Low risk	3 practices lost to follow-up
Selective reporting (reporting bias)	Low risk	Appropriate outcomes measured
Baseline similar?	Low risk	Table 1
No contamination?	Low risk	Separate practices
Risk of bias overall?	Unclear risk	Unable to assess

Frijiling 2003

Methods	Design: cluster RCT
Participants	Country: Netherlands Setting: primary care or outpatient Specialty: GP/family physician N health professionals: 185 N patients: -
Interventions	Description of Groups: feedback + outreach vs control
Outcomes	Targeted behaviour: % compliance with guidelines for cardiovascular risk mgmt Baseline performance: unclear
Notes	Format: verbal and written Source: investigators Frequency: more than monthly Instructions: action plan provided, but no specific target Nature of change: mix of increase and decrease or change behaviours

Frijiling 2003 *(Continued)*

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Quote "randomly allocated...random number generator"
Allocation concealment (selection bias)	Low risk	Quote "person responsible for the randomization process was blind to identities"
Blinding (performance bias and detection bias) All outcomes	Unclear risk	Physician report data
Incomplete outcome data (attrition bias) All outcomes	Low risk	3 of 124 practices dropped out
Selective reporting (reporting bias)	Low risk	Appropriate outcomes measured
Baseline similar?	Low risk	See Table 1
No contamination?	Low risk	Separate practices
Risk of bias overall?	Unclear risk	Unable to assess

Gama 1992

Methods	Design: cluster RCT
Participants	Country: UK Setting: inpatient Specialty: internists N health professionals: 5 N patients: 4376
Interventions	Description of Groups: AF vs control
Outcomes	Targeted behaviour: lab use Baseline performance: unclear
Notes	Format: written Source: unclear Frequency: monthly Instructions: no explicit target or action plan Nature of change: decrease

Gama 1992 *(Continued)*

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Unable to assess
Allocation concealment (selection bias)	Unclear risk	Unable to assess
Blinding (performance bias and detection bias) All outcomes	Unclear risk	Unable to assess
Incomplete outcome data (attrition bias) All outcomes	Low risk	5 of 5 physicians stayed in trial
Selective reporting (reporting bias)	Unclear risk	Unable to assess
Baseline similar?	High risk	See Table 1
No contamination?	Unclear risk	Unable to assess
Risk of bias overall?	High risk	Baseline dissimilar, many others unclear, possibly not randomized

Gehlbach 1984

Methods	Design: patient or provider level RCT
Participants	Country: USA Setting: primary care or outpatient Specialty: GP/family physicians N health professionals: 32 N patients: -
Interventions	Description of Groups: AF vs control
Outcomes	Targeted behaviour: % generic prescriptions Baseline performance: low
Notes	Format: written Source: unclear Frequency: monthly Instructions: no explicit target or action plan Nature of change: increase

Gehlbach 1984 *(Continued)*

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Unable to assess
Allocation concealment (selection bias)	Low risk	Cluster trial, allocation after recruitment completed
Blinding (performance bias and detection bias) All outcomes	Unclear risk	Data collected from medication log sheet
Incomplete outcome data (attrition bias) All outcomes	Low risk	All physicians stayed in study
Selective reporting (reporting bias)	High risk	Drugs prescribed unclear
Baseline similar?	Low risk	See Figure 1
No contamination?	Unclear risk	Unable to assess
Risk of bias overall?	High risk	Selective reporting, many unclear

Goff 2003

Methods	Design: cluster RCT
Participants	Country: USA Setting: primary care or outpatient Specialty: GP/family physicians, internists N health professionals: 605 N patients: 1570
Interventions	Description of Groups: AF + reminders vs control
Outcomes	Targeted behaviour: % compliance with guidelines for cardiovascular prescribing Baseline performance: unclear
Notes	Format: written Source: investigators Frequency: less than monthly, more than once Instructions: explicit, measurable target but no action plan Nature of change: increase

Goff 2003 *(Continued)*

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Unable to assess
Allocation concealment (selection bias)	Low risk	Cluster trial, allocation after recruitment completed
Blinding (performance bias and detection bias) All outcomes	Low risk	Administrative databases used for data collection
Incomplete outcome data (attrition bias) All outcomes	Low risk	Similar numbers of dropouts in both groups due to closings
Selective reporting (reporting bias)	Low risk	Appropriate outcomes measured
Baseline similar?	Low risk	Patient groups comparable at baseline
No contamination?	Low risk	Practices separate
Risk of bias overall?	Low risk	

Grady 1997

Methods	Design: cluster RCT
Participants	Country: USA Setting: primary care or outpatient Specialty: GP/family physicians, internists N health professionals: 95 N patients: 11 426
Interventions	Description of Groups: usual care vs physician reminders vs feedback + reminders + incentives
Outcomes	Targeted behaviour: mammography referrral, completion and compliance rates Baseline performance: low
Notes	Format: written Source: unclear Frequency: less than monthly, more than once Instructions: no explicit target or action plan Nature of change: increase

Grady 1997 *(Continued)*

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Computer generated randomization
Allocation concealment (selection bias)	Unclear risk	Unable to assess
Blinding (performance bias and detection bias) All outcomes	Unclear risk	Unable to assess
Incomplete outcome data (attrition bias) All outcomes	Low risk	87% followup of physicians
Selective reporting (reporting bias)	Low risk	Appropriate outcomes measured
Baseline similar?	Unclear risk	Unable to assess
No contamination?	Low risk	Separate practices
Risk of bias overall?	Unclear risk	Unable to assess

Guadagnoli 2000

Methods	Design: cluster RCT
Participants	Country: USA Setting: inpatient Specialty: surgeons (and entire hospital) N health professionals: 28 hospitals N patients: 1264
Interventions	Description of Groups: AF vs AF + opinion leader + education
Outcomes	Targeted behaviour: breast conserving surgery Baseline performance: moderate
Notes	Format: written Source: investigators Frequency: once only Instructions: no explicit target or action plan Nature of change: mix of increase and decrease or change behaviours

Guadagnoli 2000 *(Continued)*

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Unable to assess
Allocation concealment (selection bias)	Low risk	Cluster trial, allocation after recruitment completed
Blinding (performance bias and detection bias) All outcomes	Unclear risk	Unable to assess
Incomplete outcome data (attrition bias) All outcomes	Low risk	All hospitals with at least 7 cases participated in follow-up
Selective reporting (reporting bias)	Low risk	Appropriate outcomes measured
Baseline similar?	Unclear risk	Unable to assess
No contamination?	Low risk	Separate hospitals
Risk of bias overall?	Unclear risk	Unable to assess

Gullion 1988

Methods	Design: patient or provider level RCT
Participants	Country: USA Setting: primary care or outpatient Specialty: GP/family physicians N health professionals: 111 N patients: 2044
Interventions	Description of Groups: usual care vs feedback from medical records + education vs feedback from patient surveys+education vs both
Outcomes	Targeted behaviour: % patients with controlled blood pressure Baseline performance: unclear
Notes	Format: verbal and written Source: supervisor Frequency: once only Instructions: no explicit target or action plan Nature of change: mix of increase and decrease or change behaviours

Gullion 1988 *(Continued)*

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Quote "Stratified random assignment"
Allocation concealment (selection bias)	Unclear risk	Unable to assess
Blinding (performance bias and detection bias) All outcomes	Low risk	Quote "medical abstractors, blinded to conditions"
Incomplete outcome data (attrition bias) All outcomes	Low risk	Minimal number (5/111) lost to followup and at least one from all 4 groups
Selective reporting (reporting bias)	Low risk	Appropriate outcomes measured
Baseline similar?	Low risk	Patients comparable at baseline, see Table 3
No contamination?	Unclear risk	Unable to assess
Risk of bias overall?	Unclear risk	Unable to assess

Hayes 2001

Methods	Design: cluster RCT
Participants	Country: USA Setting: inpatient Specialty: internists (and entire hospital) N health professionals: 29 hospitals N patients: -
Interventions	Description of Groups: AF vs AF + opinion leader + QI
Outcomes	Targeted behaviour: rates of achieving a quality indicator for VTE Baseline performance: unclear
Notes	Format: written Source: unclear Frequency: once only Instructions: action plan provided, but no specific target Nature of change: increase

Hayes 2001 *(Continued)*

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Random number table
Allocation concealment (selection bias)	Low risk	Cluster trial, allocation after recruitment completed
Blinding (performance bias and detection bias) All outcomes	Low risk	Quote "Abstractors were not informed of the hospital's intervention status."
Incomplete outcome data (attrition bias) All outcomes	Low risk	Only one of 29 hospitals dropped out
Selective reporting (reporting bias)	Low risk	Appropriate outcomes measured
Baseline similar?	Low risk	See Table 1
No contamination?	Low risk	Separate hospitals
Risk of bias overall?	Low risk	

Hayes 2002

Methods	Design: cluster RCT
Participants	Country: USA Setting: inpatient Specialty: internists N health professionals: unclear, (32 hospitals) N patients: 2365
Interventions	Description of Groups: AF vs AF with educational outreach by opinion leader
Outcomes	Targeted behaviour: quality indicators for heart failure (4)
Notes	Format: written Source: unclear Frequency: once only Instructions: action plan but no specific target Nature of change: increase

Risk of bias

Hayes 2002 *(Continued)*

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Unable to assess
Allocation concealment (selection bias)	Low risk	Cluster trial, allocation after recruitment completed
Blinding (performance bias and detection bias) All outcomes	Unclear risk	Unable to assess
Incomplete outcome data (attrition bias) All outcomes	Low risk	Minimal number lost to follow-up
Selective reporting (reporting bias)	Low risk	Appropriate outcomes measured
Baseline similar?	Unclear risk	Unable to assess
No contamination?	Low risk	Separate hospitals
Risk of bias overall?	Unclear risk	Unable to assess

Heller 2001

Methods	Design: cluster RCT
Participants	Country: Austraila Setting: inpatient Specialty: internists (and entire hospital) N health professionals: 37 hospitals N patients: -
Interventions	Description of Groups: feedback + education vs control
Outcomes	Targeted behaviour: % compliance with guidelines for angina Baseline performance: unclear
Notes	Format: verbal Source: supervisor Frequency: once only Instructions: no explicit target or action plan Nature of change: mix of increase and decrease or change behaviours
Risk of bias	

Heller 2001 *(Continued)*

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Unable to assess
Allocation concealment (selection bias)	Unclear risk	Unable to assess
Blinding (performance bias and detection bias) All outcomes	Unclear risk	Unable to assess
Incomplete outcome data (attrition bias) All outcomes	Low risk	Approximatley 10% lost to follow-up for behavioural outcomes (Box 2)
Selective reporting (reporting bias)	Low risk	Appropriate outcomes measured
Baseline similar?	Unclear risk	Unable to assess
No contamination?	Low risk	Separate hospitals
Risk of bias overall?	Unclear risk	Unable to assess

Hemminiki 1992

Methods	Design: cluster RCT
Participants	Country: Finland Setting: inpatient Specialty: obstetricians (and nurses) N health professionals: 53 hospitals N patients: -
Interventions	Description of Groups: AF vs control
Outcomes	Targeted behaviour: % vaginal deliveries Baseline performance: unclear
Notes	Format: written Source: rep from employer or quality assurance org Frequency: once only Instructions: no explicit target or action plan Nature of change: mix of increase and decrease or change behaviours
Risk of bias	

Hemminiki 1992 *(Continued)*

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Unable to assess
Allocation concealment (selection bias)	Low risk	Cluster trial, allocation after recruitment completed
Blinding (performance bias and detection bias) All outcomes	Low risk	Data collected from registers
Incomplete outcome data (attrition bias) All outcomes	Low risk	Data on 48 of 52 hospitals
Selective reporting (reporting bias)	Low risk	Appropriate outcomes measured
Baseline similar?	Unclear risk	Unable to assess
No contamination?	Low risk	Separate hospitals
Risk of bias overall?	Unclear risk	Unable to assess

Hendryx 1998

Methods	Design: cluster RCT
Participants	Country: USA Setting: inpatient Specialty: internists (and nurses) N health professionals: 20 hospitals N patients: -
Interventions	Description of Groups: feedback + outreach + education + telephone consult service vs control
Outcomes	Targeted behaviour: compliance with ICU guidelines Baseline performance: unclear
Notes	Format: verbal and written Source: supervisor Frequency: once only Instructions: action plan provided, but no specific target Nature of change: mix of increase and decrease or change behaviours
Risk of bias	

Hendryx 1998 *(Continued)*

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Unable to assess
Allocation concealment (selection bias)	Low risk	Cluster trial, allocation after recruitment completed
Blinding (performance bias and detection bias) All outcomes	Unclear risk	Unable to assess
Incomplete outcome data (attrition bias) All outcomes	Low risk	All hospitals followed up
Selective reporting (reporting bias)	Low risk	Appropriate outcomes measured
Baseline similar?	Unclear risk	Unable to assess
No contamination?	Low risk	Separate hospitals
Risk of bias overall?	Unclear risk	Unable to assess

Herbert 2004

Methods	Design: cluster RCT
Participants	Country: Canada Setting: outpatient Specialty: gp/family medicine N health professionals: 200 N patients: 3128
Interventions	Description of Groups: usual care vs feedback vs education vs feedback plus education
Outcomes	Targeted behaviour: use of thiazide as first antihypertensive Baseline performance: low
Notes	Format: written Source: investigators Frequency: once only Instructions: action plan provided but no specific target Nature of change: increase
Risk of bias	

Herbert 2004 *(Continued)*

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Appears random with matching
Allocation concealment (selection bias)	Low risk	Cluster trial, allocation after recruitment completed
Blinding (performance bias and detection bias) All outcomes	Low risk	Arbitrary codes used for labeling
Incomplete outcome data (attrition bias) All outcomes	Low risk	See Figure 1
Selective reporting (reporting bias)	Unclear risk	Unable to assess
Baseline similar?	Unclear risk	Unable to assess
No contamination?	Low risk	Separate hospitals
Risk of bias overall?	Low risk	

Herrin 2006

Methods	Design: cluster RCT
Participants	Country: USA Setting: primary care or outpatient Specialty: GP/family physicians, internists N health professionals: 92 N patients: 2155
Interventions	Description of Groups: aggregate feedback vs patient specific feedback vs patient specific feedback plus diabetes nurse (case mgmt)
Outcomes	Targeted behaviour: diabetes management Baseline performance: unclear
Notes	Format: unclear Source: rep from employer or quality assurance org Frequency: once only Instructions: no explicit target or action plan Nature of change: increase
Risk of bias	

Herrin 2006 *(Continued)*

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Unable to assess
Allocation concealment (selection bias)	Low risk	Performed on all units at start
Blinding (performance bias and detection bias) All outcomes	Unclear risk	Unable to assess
Incomplete outcome data (attrition bias) All outcomes	Low risk	See Figure
Selective reporting (reporting bias)	Low risk	Appropriate outcomes measured
Baseline similar?	Low risk	See Table 1
No contamination?	Low risk	Separate practices
Risk of bias overall?	Unclear risk	Unable to assess

Hershey 1986

Methods	Design: cluster RCT
Participants	Country: USA Setting: inpatient Specialty: internists N health professionals: 48 N patients: -
Interventions	Description of Groups: AF vs control
Outcomes	Targeted behaviour: number of prescriptions per patient Baseline performance: unclear
Notes	Format: written Source: unclear Frequency: monthly Instructions: action plan provided, but no specific target Nature of change: decrease
Risk of bias	

Hershey 1986 *(Continued)*

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Unable to assess
Allocation concealment (selection bias)	Unclear risk	Unable to assess
Blinding (performance bias and detection bias) All outcomes	Low risk	Data centrally computer generated
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Unable to assess
Selective reporting (reporting bias)	Low risk	Appropriate outcomes measured
Baseline similar?	Unclear risk	Unable to assess
No contamination?	Low risk	Separate firms
Risk of bias overall?	Unclear risk	Unable to assess

Hershey 1988

Methods	Design: cluster RCT
Participants	Country: USA Setting: inpatient Specialty: internists N health professionals: 50, (4 practices) N patients: 3000
Interventions	Description of Groups: AF vs AF plus written education
Outcomes	Targeted behaviour: number of prescriptions Baseline performance: unclear
Notes	Format: written Source: investigators Frequency: monthly Instructions: no action plan or explicit target Nature of change: mix / unclear
Risk of bias	

Hershey 1988 *(Continued)*

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Unable to assess
Allocation concealment (selection bias)	Low risk	Cluster trial, allocation after recruitment completed
Blinding (performance bias and detection bias) All outcomes	Low risk	Computerized data reported
Incomplete outcome data (attrition bias) All outcomes	Low risk	All firms followed up
Selective reporting (reporting bias)	Low risk	Appropriate outcomes measured
Baseline similar?	Low risk	Table 1
No contamination?	Unclear risk	Unable to assess
Risk of bias overall?	Unclear risk	Unable to assess

Hillman 1998

Methods	Design: cluster RCT
Participants	Country: USA Setting: primary care or outpatient Specialty: GP/family physicians, internists N health professionals: 52 practices N patients: -
Interventions	Description of Groups: AF + incentive vs control
Outcomes	Targeted behaviour: cancer screening Baseline performance: low
Notes	Format: written Source: unclear Frequency: less than monthly, more than once Instructions: no explicit target or action plan Nature of change: increase
Risk of bias	

Hillman 1998 *(Continued)*

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Stratified randomization
Allocation concealment (selection bias)	Low risk	Cluster trial, allocation after recruitment completed
Blinding (performance bias and detection bias) All outcomes	Unclear risk	Unable to assess
Incomplete outcome data (attrition bias) All outcomes	Low risk	Data collected from all sites
Selective reporting (reporting bias)	Low risk	Appropriate outcomes measured
Baseline similar?	Unclear risk	Unable to assess
No contamination?	Low risk	Separate sites
Risk of bias overall?	Unclear risk	Unable to assess

Hillman 1999

Methods	Design: cluster RCT
Participants	Country: USA Setting: primary care or outpatient Specialty: pediatricians, GP/family physicians N health professionals: 49 practices N patients: -
Interventions	Description of Groups: usual care vs feedback vs feedback + incentive
Outcomes	Targeted behaviour: compliance with well child care guidelines Baseline performance: unclear
Notes	Format: written Source: unclear Frequency: less than monthly, more than once Instructions: no explicit target or action plan Nature of change: increase
Risk of bias	

Hillman 1999 *(Continued)*

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Stratified, randomized
Allocation concealment (selection bias)	Low risk	Cluster trial, allocation after recruitment completed
Blinding (performance bias and detection bias) All outcomes	Low risk	Reviewers blinded to intervention sites
Incomplete outcome data (attrition bias) All outcomes	Low risk	49 of 53 sites completed study
Selective reporting (reporting bias)	Low risk	Appropriate outcomes measured
Baseline similar?	Low risk	See Table 2
No contamination?	Low risk	Separate sites
Risk of bias overall?	Low risk	

Holm 1990

Methods	Design: cluster RCT
Participants	Country: Denmark Setting: primary care or outpatient Specialty: GP/family physicians N health professionals: 365 N patients: -
Interventions	Description of Groups: usual care vs education vs feedback + education
Outcomes	Targeted behaviour: prescribing of benzodiazepines Baseline performance: unclear
Notes	Format: written Source: unclear Frequency: once only Instructions: no explicit target or action plan Nature of change: decrease
Risk of bias	

Holm 1990 *(Continued)*

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	"Drawing lots" used as method of randomization
Allocation concealment (selection bias)	Low risk	Cluster trial, allocation after recruitment completed
Blinding (performance bias and detection bias) All outcomes	Unclear risk	Unable to assess
Incomplete outcome data (attrition bias) All outcomes	Low risk	Minimal number of dropouts
Selective reporting (reporting bias)	Low risk	Appropriate outcomes measured
Baseline similar?	Low risk	See Table 1
No contamination?	Low risk	Separate practices
Risk of bias overall?	Unclear risk	Unable to assess

Hux 1999

Methods	Design: cluster RCT
Participants	Country: Canada Setting: primary care or outpatient Specialty: GP/family physicians N health professionals: 251 N patients: -
Interventions	Description of Groups: AF vs control
Outcomes	Targeted behaviour: appropriate antibiotic prescribing Baseline performance: low
Notes	Format: written Source: unclear Frequency: less than monthly, more than once Instructions: no explicit target or action plan Nature of change: mix of increase and decrease or change behaviours
Risk of bias	

Hux 1999 *(Continued)*

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Unable to assess
Allocation concealment (selection bias)	Low risk	Cluster trial, allocation after recruitment completed
Blinding (performance bias and detection bias) All outcomes	Low risk	Unique identifiers used
Incomplete outcome data (attrition bias) All outcomes	Low risk	Data avaiulable for 250/251 physicians
Selective reporting (reporting bias)	Low risk	Appropriate outcomes measured
Baseline similar?	Low risk	See Table 1
No contamination?	Low risk	Physicans recruited from separate addresses
Risk of bias overall?	Low risk	

Kahan 2009

Methods	Design: patient or provider level RCT
Participants	Country: Israel Setting: primary care or outpatient Specialty: GP/family physicians, internists N health professionals: 298 N patients: -
Interventions	Description of Groups: usual care vs feedback vs seminar vs feedback plus seminar
Outcomes	Targeted behaviour: antibiotic prescribing for UTI Baseline performance: low
Notes	Format: written Source: supervisor Frequency: once only Instructions: no explicit target or action plan Nature of change: mix of increase and decrease or change behaviours
Risk of bias	

Kahan 2009 *(Continued)*

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Unable to assess
Allocation concealment (selection bias)	Low risk	Performed on all units at start
Blinding (performance bias and detection bias) All outcomes	Unclear risk	Unable to assess
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Unable to assess
Selective reporting (reporting bias)	Low risk	Appropriate outcomes measured
Baseline similar?	Low risk	Table 4
No contamination?	Unclear risk	Unable to assess
Risk of bias overall?	Unclear risk	Unable to assess

Kerry 2000

Methods	Design: cluster RCT
Participants	Country: UK Setting: primary care or outpatient Specialty: GP/family physicians N health professionals: 175 N patients: -
Interventions	Description of Groups: AF vs control
Outcomes	Targeted behaviour: xray referral rates Baseline performance: unclear
Notes	Format: written Source: unclear Frequency: once only Instructions: no explicit target or action plan Nature of change: decrease
Risk of bias	

Kerry 2000 *(Continued)*

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Stratified randomization used
Allocation concealment (selection bias)	Low risk	Cluster trial, allocation after recruitment completed
Blinding (performance bias and detection bias) All outcomes	Low risk	Data collected from central computer registry
Incomplete outcome data (attrition bias) All outcomes	Low risk	Data collected on all practices
Selective reporting (reporting bias)	Low risk	Appropriate outcomes measured
Baseline similar?	Low risk	Table 1
No contamination?	Low risk	Practices separate
Risk of bias overall?	Low risk	

Kiefe 2001

Methods	Design: patient or provider level RCT
Participants	Country: USA Setting: primary care or outpatient Specialty: GP/family physicians, internists N health professionals: 97 N patients: 2978
Interventions	Description of Groups: usual care vs AF vs AF with benchmark
Outcomes	Targeted behaviour: quality indicators for prevention (5) Baseline performance: unclear
Notes	Format: verbal and written Source: rep from employer or quality assurance org Frequency: less than monthly, more than once Instructions: explicit, measurable target but no action plan Nature of change: increase
Risk of bias	

Kiefe 2001 *(Continued)*

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Unable to assess
Allocation concealment (selection bias)	Low risk	Cluster trial, allocation after recruitment completed
Blinding (performance bias and detection bias) All outcomes	Unclear risk	Unable to assess
Incomplete outcome data (attrition bias) All outcomes	Low risk	Equal numbers in both groups lost to follow up (Figure 1)
Selective reporting (reporting bias)	Low risk	Appropriate outcomes measured
Baseline similar?	Unclear risk	Unable to assess
No contamination?	Unclear risk	Unable to assess
Risk of bias overall?	Unclear risk	Unable to assess

Kim 1999

Methods	Design: cluster RCT
Participants	Country: Scotland Setting: primary care or outpatient Specialty: GP/family physicians, internists N health professionals: 48 N patients: 1810
Interventions	Description of Groups: AF + educational outreach vs control
Outcomes	Targeted behaviour: advice about preventive services Baseline performance: unclear
Notes	Format: verbal and written Source: unclear Frequency: less than monthly, more than once Instructions: action plan provided, but no specific target Nature of change: increase
Risk of bias	

Kim 1999 *(Continued)*

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Unable to assess
Allocation concealment (selection bias)	Low risk	Cluster trial, allocation after recruitment completed
Blinding (performance bias and detection bias) All outcomes	Unclear risk	Unable to assess
Incomplete outcome data (attrition bias) All outcomes	High risk	Only reviewed charts of patients who responded
Selective reporting (reporting bias)	Unclear risk	Unable to assess
Baseline similar?	Unclear risk	Unable to assess
No contamination?	Unclear risk	Unable to assess
Risk of bias overall?	High risk	Incomplete follow up

Kinsinger 1998

Methods	Design: cluster RCT
Participants	Country: USA Setting: primary care or outpatient Specialty: GP/family physicians, internists N health professionals: 62 practices N patients: 2874
Interventions	Description of Groups: AF with office QI support vs AF
Outcomes	Targeted behaviour: screening rates breast cancer Baseline performance: moderate
Notes	Format: verbal and written Source: investigators Frequency: once only Instructions: action plan provided, but no specific target Nature of change: increase
Risk of bias	

Kinsinger 1998 *(Continued)*

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Randomization conducted by statistitian
Allocation concealment (selection bias)	Low risk	Cluster trial, allocation after recruitment completed
Blinding (performance bias and detection bias) All outcomes	Low risk	Research assistants who collected data were blinded
Incomplete outcome data (attrition bias) All outcomes	Low risk	58/62 practices provided data
Selective reporting (reporting bias)	Low risk	Appropriate outcomes measured
Baseline similar?	Low risk	Table 1 and Table 2
No contamination?	Low risk	Separate practices
Risk of bias overall?	Low risk	

Kogan 2003

Methods	Design: cluster RCT
Participants	Country: USA Setting: inpatient Specialty: internists N health professionals: 44 N patients: -
Interventions	Description of Groups: AF vs control
Outcomes	Targeted behaviour: total performance scores (% of indicated action taken) for prevetive health and disease management Baseline performance: unclear
Notes	Format: written Source: supervisor Frequency: once only Instructions: no explicit target or action plan Nature of change: increase
Risk of bias	

Kogan 2003 *(Continued)*

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Unable to assess
Allocation concealment (selection bias)	Low risk	Cluster trial, allocation after recruitment completed
Blinding (performance bias and detection bias) All outcomes	Low risk	Abstractors blinded
Incomplete outcome data (attrition bias) All outcomes	Low risk	Fewer patients than planned for but good follow-up
Selective reporting (reporting bias)	Low risk	Appropriate outcomes measured
Baseline similar?	Low risk	Table 1 and 2
No contamination?	Unclear risk	Unable to assess
Risk of bias overall?	Unclear risk	Unable to assess

Kritchevsky 2008

Methods	Design: cluster RCT
Participants	Country: USA Setting: inpatient Specialty: surgery N health professionals: 44 hospitals N patients: 8800
Interventions	Description of Groups: feedback vs feedback plus learning collaboratives
Outcomes	Targeted behaviour: antiobiotics pre-surgery Baseline performance: high
Notes	Format: unclear Source: investigators Frequency: once only Instructions: no explicit target or action plan Nature of change: increase
Risk of bias	

Kritchevsky 2008 *(Continued)*

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Paired, blinded randomization
Allocation concealment (selection bias)	Low risk	Cluster trial, allocation after recruitment completed
Blinding (performance bias and detection bias) All outcomes	Low risk	Authors report high reliability of main outcomes so judged to be low risk
Incomplete outcome data (attrition bias) All outcomes	Low risk	All hospitals provided data
Selective reporting (reporting bias)	Low risk	Appropriate outcomes measured
Baseline similar?	Low risk	Table 2
No contamination?	Low risk	Separate hospitals
Risk of bias overall?	Low risk	

Lagerløv 2000

Methods	Design: cluster RCT
Participants	Country: Norway Setting: primary care or outpatient Specialty: GP/family physician N health professionals: 199, (32 communities) N patients: unclear
Interventions	Description of Groups: usual care vs feedback plus education
Outcomes	Targeted behaviour: appropriate prescribing for asthma, UTI Baseline performance: low
Notes	Format: verbal and written Source: investigators Frequency: once only Instructions: goal-setting and action plans Nature of change: mix / unclear
Risk of bias	

Lagerløv 2000 *(Continued)*

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Unable to assess
Allocation concealment (selection bias)	Unclear risk	Unable to assess
Blinding (performance bias and detection bias) All outcomes	Unclear risk	Unable to assess
Incomplete outcome data (attrition bias) All outcomes	Low risk	Data provided on all physicians
Selective reporting (reporting bias)	Low risk	Appropriate outcomes measured
Baseline similar?	Low risk	Table 3
No contamination?	Low risk	GPs separate
Risk of bias overall?	Unclear risk	Unable to assess

Laskshminarayan 2010

Methods	Design: cluster RCT
Participants	Country: USA Setting: inpatient Specialty: GP/family physician, internists N health professionals: unclear (19 hospitals) N patients: 2305
Interventions	Description of Groups: AF vs AF plus educational outreach from opinion leaders and continuous quality improvement
Outcomes	Targeted behaviour: compliance with guidelines for stroke (10 outcomes) Baseline performance: moderate
Notes	Format: verbal and written Source: opinion leaders / respected senior colleagues Frequency: once only Instructions: no explicit target or action plan Nature of change: increase
Risk of bias	

Laskshminarayan 2010 *(Continued)*

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Unable to assess
Allocation concealment (selection bias)	Low risk	Cluster trial, allocation after recruitment completed
Blinding (performance bias and detection bias) All outcomes	Low risk	Abstractors blinded
Incomplete outcome data (attrition bias) All outcomes	Low risk	See Figure 1
Selective reporting (reporting bias)	Low risk	Appropriate outcomes measured
Baseline similar?	Unclear risk	Unable to assess
No contamination?	Low risk	Separate hospitals
Risk of bias overall?	Unclear risk	Unable to assess

Linn 1980

Methods	Design: cluster RCT
Participants	Country: USA Setting: inpatient Specialty: ER N health professionals: 298 N patients: 2664
Interventions	Description of Groups: AF+education+hotline vs control
Outcomes	Targeted behaviour: deviations from guidelines for burn care in er Baseline performance: unclear
Notes	Format: written Source: unclear Frequency: unclear Instructions: action plan provided, but no specific target Nature of change: increase
Risk of bias	

Linn 1980 *(Continued)*

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Random number table
Allocation concealment (selection bias)	Low risk	Cluster trial, allocation after recruitment completed
Blinding (performance bias and detection bias) All outcomes	Unclear risk	Unable to assess
Incomplete outcome data (attrition bias) All outcomes	Low risk	No attrition of primary outcome (process of care)
Selective reporting (reporting bias)	Low risk	Appropriate outcomes measured
Baseline similar?	Unclear risk	Unable to assess
No contamination?	Low risk	Separate hospitals
Risk of bias overall?	Unclear risk	Unable to assess

Lobach 1996

Methods	Design: patient or provider level RCT
Participants	Country: USA Setting: primary care or outpatient Specialty: GP/family physicians N health professionals: 45 N patients: -
Interventions	Description of Groups: AF vs control
Outcomes	Targeted behaviour: compliance with guidelines for diabetes Baseline performance: unclear
Notes	Format: written Source: unclear Frequency: more than monthly Instructions: no explicit target or action plan Nature of change: increase
Risk of bias	

Lobach 1996 *(Continued)*

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Unable to assess
Allocation concealment (selection bias)	Low risk	Cluster trial, allocation after recruitment completed
Blinding (performance bias and detection bias) All outcomes	Low risk	Outcomes from computerized data
Incomplete outcome data (attrition bias) All outcomes	Low risk	No clinicians lost to follow-up
Selective reporting (reporting bias)	Unclear risk	Unable to assess
Baseline similar?	Unclear risk	Unable to assess
No contamination?	Unclear risk	Unable to assess
Risk of bias overall?	Unclear risk	Unable to assess

Lomas 1991

Methods	Design: cluster RCT
Participants	Country: Canada Setting: inpatient Specialty: obgyn N health professionals: 76 N patients: -
Interventions	Description of Groups: feedback + education vs opinion leaders + education vs control
Outcomes	Targeted behaviour: quality indicators for labour and delivery Baseline performance: unclear
Notes	Format: verbal and written Source: investigators Frequency: less than monthly, more than once Instructions: no explicit target or action plan Nature of change: decrease
Risk of bias	

Lomas 1991 *(Continued)*

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Unable to assess
Allocation concealment (selection bias)	Low risk	Cluster trial, allocation after recruitment completed
Blinding (performance bias and detection bias) All outcomes	Low risk	Chart audits done by trained staff
Incomplete outcome data (attrition bias) All outcomes	Low risk	Data on all hospitals provided
Selective reporting (reporting bias)	Low risk	Appropriate outcomes measured
Baseline similar?	Low risk	Table 1
No contamination?	Low risk	Separate hospitals
Risk of bias overall?	Low risk	

Mainous 2000

Methods	Design: cluster RCT
Participants	Country: USA Setting: primary care or outpatient Specialty: GP/family physicians, pediatricians, internists N health professionals: 216 N patients: -
Interventions	Description of Groups: usual care vs feedback vs patient education vs both
Outcomes	Targeted behaviour: mean antibiotic prescribing rates Baseline performance: high
Notes	Format: written Source: unclear Frequency: once only Instructions: no explicit target or action plan Nature of change: decrease
Risk of bias	

Mainous 2000 *(Continued)*

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Unable to assess
Allocation concealment (selection bias)	Unclear risk	Unable to assess
Blinding (performance bias and detection bias) All outcomes	Low risk	Data collected from Medicaid drug records
Incomplete outcome data (attrition bias) All outcomes	Low risk	Data from all physicians provided
Selective reporting (reporting bias)	Low risk	Appropriate outcomes measured
Baseline similar?	Low risk	Table 2
No contamination?	Low risk	Physicians were in separate practices
Risk of bias overall?	Unclear risk	Unable to assess

Martin 1980

Methods	Design: cluster RCT
Participants	Country: USA Setting: inpatient Specialty: internists N health professionals: 24 N patients: -
Interventions	Description of Groups: AF vs incentives vs control
Outcomes	Targeted behaviour: mean tests per patient admission Baseline performance: unclear
Notes	Format: verbal Source: investigators Frequency: more than monthly Instructions: action plan provided, but no specific target Nature of change: decrease
Risk of bias	

Martin 1980 *(Continued)*

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Unable to assess
Allocation concealment (selection bias)	Low risk	Cluster trial, allocation after recruitment completed
Blinding (performance bias and detection bias) All outcomes	Unclear risk	Unable to assess
Incomplete outcome data (attrition bias) All outcomes	Low risk	No attrition on process of care outcomes
Selective reporting (reporting bias)	Low risk	Appropriate outcomes measured
Baseline similar?	Low risk	Table 1
No contamination?	Unclear risk	Unable to assess
Risk of bias overall?	Unclear risk	Unable to assess

Marton 1985

Methods	Design: patient or provider level RCT
Participants	Country: USA Setting: primary care or outpatient Specialty: internists N health professionals: 57 N patients: -
Interventions	Description of Groups: usual care vs AF vs education vs both
Outcomes	Targeted behaviour: lab tests per patient admission Baseline performance: unclear
Notes	Format: written Source: unclear Frequency: monthly Instructions: no explicit target or action plan Nature of change: decrease
Risk of bias	

Marton 1985 *(Continued)*

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Unable to assess
Allocation concealment (selection bias)	Low risk	Cluster trial, allocation after recruitment completed
Blinding (performance bias and detection bias) All outcomes	Unclear risk	Unable to assess
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	No description of whether or not subjects lost to follow-up.
Selective reporting (reporting bias)	Low risk	Appropriate outcomes measured
Baseline similar?	Low risk	See Table 1
No contamination?	Unclear risk	Unable to assess
Risk of bias overall?	Unclear risk	Unable to assess

Mayer 1998

Methods	Design: cluster RCT
Participants	Country: USA Setting: primary care or outpatient Specialty: pharmacists N health professionals: 138 N patients: -
Interventions	Description of Groups: feedback + education + reminders + incentives vs control
Outcomes	Targeted behaviour: skin cancer prevention advice Baseline performance: low
Notes	Format: written Source: unclear Frequency: more than monthly Instructions: no explicit target or action plan Nature of change: increase
Risk of bias	

Mayer 1998 *(Continued)*

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Unable to assess
Allocation concealment (selection bias)	Low risk	Cluster trial, allocation after recruitment completed
Blinding (performance bias and detection bias) All outcomes	Low risk	Quote "Confederates blinded to pharmacy study conditions"
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Unable to assess
Selective reporting (reporting bias)	Low risk	Appropriate outcomes measured
Baseline similar?	Low risk	Groups similar
No contamination?	Low risk	Pharmacies separate
Risk of bias overall?	Unclear risk	Unable to assess

McAlister 1986

Methods	Design: cluster RCT
Participants	Country: Canada Setting: primary care or outpatient Specialty: GP/family physicians N health professionals: 60 N patients: -
Interventions	Description of Groups: feedback + patient reminders vs control
Outcomes	Targeted behaviour: management of hypertension Baseline performance: unclear
Notes	Format: written Source: unclear Frequency: unclear Instructions: action plan provided, but no specific target Nature of change: mix of increase and decrease or change behaviours
Risk of bias	

McAlister 1986 *(Continued)*

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Shuffled cards
Allocation concealment (selection bias)	Low risk	Cluster trial, allocation after recruitment completed
Blinding (performance bias and detection bias) All outcomes	Unclear risk	Unable to assess
Incomplete outcome data (attrition bias) All outcomes	Low risk	Equal numbers of dropouts in both groups
Selective reporting (reporting bias)	Unclear risk	Unable to assess
Baseline similar?	Low risk	Groups comparable at baseline
No contamination?	Low risk	Separate practices
Risk of bias overall?	Unclear risk	Unable to assess

McCartney 1997

Methods	Design: cluster RCT
Participants	Country: UK Setting: primary care or outpatient Specialty: GP/family physicians N health professionals: 28 N patients: 182 220
Interventions	Description of Groups: AF vs control
Outcomes	Targeted behaviour: % patients with CHD on aspirin Baseline performance: moderate
Notes	Format: written Source: unclear Frequency: once only Instructions: action plan provided, but no specific target Nature of change: increase
Risk of bias	

McCartney 1997 *(Continued)*

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Unable to assess
Allocation concealment (selection bias)	Low risk	Cluster trial, allocation after recruitment completed. Sealed envelopes used
Blinding (performance bias and detection bias) All outcomes	Low risk	Data collected by computer searches
Incomplete outcome data (attrition bias) All outcomes	Low risk	Data from all 28 practices
Selective reporting (reporting bias)	Low risk	Appropriate outcomes measured
Baseline similar?	Low risk	Figure 1
No contamination?	Low risk	Separate practices
Risk of bias overall?	Low risk	

McClellan 2003

Methods	Design: cluster RCT
Participants	Country: USA Setting: primary care or outpatient Specialty: GP/family physicians, internists, OB/GYN N health professionals: 477, (123 communities) N patients: 22,971
Interventions	Description of Groups: usual care vs AF plus education plus reminders
Outcomes	Targeted behaviour: diabetes management - tests and referrals Baseline performance: moderate
Notes	Format: written Source: rep from employer or quality assurance org Frequency: less than monthly, more than once Instructions: neither explicit goals or action plan Nature of change: increase
Risk of bias	

McClellan 2003 *(Continued)*

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Random number
Allocation concealment (selection bias)	Low risk	Cluster trial, allocation after recruitment completed
Blinding (performance bias and detection bias) All outcomes	Low risk	Data collected from database
Incomplete outcome data (attrition bias) All outcomes	Low risk	Similar dropouts in both groups
Selective reporting (reporting bias)	Low risk	Appropriate outcomes measured
Baseline similar?	Low risk	Table 1
No contamination?	Low risk	Separate practices
Risk of bias overall?	Low risk	

McClellan 2004

Methods	Design: cluster RCT
Participants	Country: USA Setting: dialysis centers Specialty: internists N health professionals: unclear (41 centers) N patients: 4280
Interventions	Description of Groups: AF vs AF plus educational outreach with continuous quality improvement
Outcomes	Targeted behaviour: proportion of patients with adequate dialysis Baseline performance: high
Notes	Format: unclear Source: rep from employer or quality assurance org Frequency: once only Instructions: neither goal-setting or action plan Nature of change: mix / unclear
Risk of bias	

McClellan 2004 *(Continued)*

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Random number allocation
Allocation concealment (selection bias)	Low risk	Cluster trial, allocation after recruitment completed
Blinding (performance bias and detection bias) All outcomes	Low risk	Lab values blinded
Incomplete outcome data (attrition bias) All outcomes	Low risk	Data from all centre in study
Selective reporting (reporting bias)	Low risk	Appropriate outcomes measured
Baseline similar?	Low risk	See Table 1
No contamination?	Low risk	Separate centres
Risk of bias overall?	Low risk	

McConnell 1982

Methods	Design: patient or provider level RCT
Participants	Country: USA Setting: primary care or outpatient Specialty: unclear N health professionals: 35 N patients: -
Interventions	Description of Groups: feedback + educaiton outreach vs control
Outcomes	Targeted behaviour: % continuing to prescribe tetracycline inappropriately Baseline performance: low
Notes	Format: verbal and written Source: rep from employer or quality assurance org Frequency: once only Instructions: no explicit target or action plan Nature of change: decrease
Risk of bias	

McConnell 1982 *(Continued)*

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Unable to assess
Allocation concealment (selection bias)	Low risk	Cluster trial, allocation after recruitment completed
Blinding (performance bias and detection bias) All outcomes	Low risk	Medicaid data used
Incomplete outcome data (attrition bias) All outcomes	Low risk	33/35 physicians followed up
Selective reporting (reporting bias)	Low risk	Appropriate outcomes measured
Baseline similar?	Unclear risk	Unable to assess
No contamination?	Unclear risk	Unable to assess
Risk of bias overall?	Unclear risk	Unable to assess

Millard 2008

Methods	Design: patient or provider level RCT
Participants	Country: Australia Setting: primary care or outpatient Specialty: GP/family physicians N health professionals: 14 N patients: unclear
Interventions	Description of Groups: usual care vs AF
Outcomes	Targeted behaviour: identification/diagnosis of dementia Baseline performance: unclear
Notes	Format: written Source: investigators Frequency: once only Instructions: neither goal-setting or action plans Nature of change: increase
Risk of bias	

Millard 2008 *(Continued)*

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Unable to assess
Allocation concealment (selection bias)	Low risk	Cluster trial, allocation after recruitment completed
Blinding (performance bias and detection bias) All outcomes	High risk	Practice staff undertook own data extraction
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Unable to assess
Selective reporting (reporting bias)	Unclear risk	Unable to assess
Baseline similar?	High risk	Table 1
No contamination?	Unclear risk	Unable to assess
Risk of bias overall?	High risk	High risk on audit process (variable between sites), baseline imbalances

Mitchell 2005

Methods	Design: cluster RCT
Participants	Country: Scotland Setting: primary care or outpatient Specialty: GP/family physicians N health professionals: 189 N patients: 20393
Interventions	Description of Groups: usual care v aggregate feedback v feedback with patient-specific risk scores
Outcomes	Targeted behaviour: control of hypertension Baseline performance: moderate
Notes	Format: written Source: investigators Frequency: less than monthly, more than once Instructions: no explicit target or action plan Nature of change: increase
Risk of bias	

Mitchell 2005 *(Continued)*

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Computer generated
Allocation concealment (selection bias)	Low risk	Cluster trial, allocation after recruitment completed
Blinding (performance bias and detection bias) All outcomes	Low risk	Administrative data
Incomplete outcome data (attrition bias) All outcomes	Low risk	52/54 GPs provided data
Selective reporting (reporting bias)	Low risk	Appropriate outcomes measured
Baseline similar?	Low risk	See Table 1
No contamination?	Low risk	Separate GPs
Risk of bias overall?	Low risk	

Moher 2001

Methods	Design: cluster RCT
Participants	Country: UK Setting: primary care or outpatient Specialty: GP/family physicians (and nurses) N health professionals: 21 practices N patients: 1906
Interventions	Description of Groups: AF vs AF + doctor recall vs AF +r ecall by nurse
Outcomes	Targeted behaviour: % adequate assessement of risk factors and drug therapy for patients with CHD Baseline performance: unclear
Notes	Format: verbal and written Source: investigators Frequency: once only Instructions: no explicit target or action plan Nature of change: increase
Risk of bias	

Moher 2001 *(Continued)*

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Unable to assess
Allocation concealment (selection bias)	Low risk	Cluster trial, allocation after recruitment completed
Blinding (performance bias and detection bias) All outcomes	Unclear risk	Unable to assess
Incomplete outcome data (attrition bias) All outcomes	Low risk	Data on all 21 practices
Selective reporting (reporting bias)	Low risk	Appropriate outcomes measured
Baseline similar?	Low risk	See Table 1
No contamination?	Low risk	Separate practices
Risk of bias overall?	Unclear risk	Unable to assess

Mold 2008

Methods	Design: cluster RCT
Participants	Country: USA Setting: primary care or outpatient Specialty: GP/family physicians N health professionals: 24 N patients: -
Interventions	Description of Groups: feedback alone vs feedback and academic detailing
Outcomes	Targeted behaviour: preventive services Baseline performance: moderate
Notes	Format: written Source: investigators Frequency: once only Instructions: no explicit target or action plan Nature of change: increase
Risk of bias	

Mold 2008 *(Continued)*

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Coin tosses
Allocation concealment (selection bias)	Low risk	Cluster trial, allocation after recruitment completed
Blinding (performance bias and detection bias) All outcomes	Low risk	Interviewers blinded
Incomplete outcome data (attrition bias) All outcomes	Low risk	All clinicians followed up
Selective reporting (reporting bias)	Low risk	Appropriate outcomes measured
Baseline similar?	High risk	Table 1
No contamination?	Low risk	Separate practices
Risk of bias overall?	Unclear risk	Unable to assess

Naughton 2007

Methods	Design:cluster RCT
Participants	Country: Ireland Setting: primary care or outpatient Specialty: GP/family physicians N health professionals: 109 N patients: 1796
Interventions	Description of Groups: feedback only vs feedback plus academic detailing
Outcomes	Targeted behaviour: prescription rates for pts with CV risk Baseline performance: high
Notes	Format: written Source: investigators Frequency: once only Instructions: no explicit target or action plan Nature of change: increase
Risk of bias	

Naughton 2007 *(Continued)*

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Unable to assess
Allocation concealment (selection bias)	Low risk	Cluster trial, allocation after recruitment completed
Blinding (performance bias and detection bias) All outcomes	Low risk	Administrative data used
Incomplete outcome data (attrition bias) All outcomes	Low risk	1 dropout of 98 GPs
Selective reporting (reporting bias)	Low risk	Appropriate outcomes measured
Baseline similar?	Low risk	Table 1
No contamination?	Low risk	Separate GPs
Risk of bias overall?	Low risk	

Naughton 2009

Methods	Design: cluster RCT
Participants	Country: Ireland Setting: primary care or outpatient Specialty: GP/family physicians N health professionals: 110 N patients: -
Interventions	Description of Groups: feedback only vs feedback plus academic detailing
Outcomes	Targeted behaviour: prescribing rates of antibiotics Baseline performance: unclear
Notes	Format: written Source: investigators Frequency: once only Instructions: no explicit target or action plan Nature of change: decrease
Risk of bias	

Naughton 2009 *(Continued)*

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Unable to assess
Allocation concealment (selection bias)	Low risk	Cluster trial, allocation after recruitment completed
Blinding (performance bias and detection bias) All outcomes	Low risk	Administrative data
Incomplete outcome data (attrition bias) All outcomes	Low risk	Data on all GPs provided
Selective reporting (reporting bias)	Low risk	Appropriate outcomes measured
Baseline similar?	Low risk	See Table 1
No contamination?	Low risk	GPs in separate practices
Risk of bias overall?	Low risk	

Nilsson 2001

Methods	Design: cluster RCT
Participants	Country: Sweden Setting: primary care or outpatient Specialty: GP/family physicians N health professionals: 40 N patients: 45982
Interventions	Description of Groups: feedback + opinion leaders + educational outreach vs control (with active controls x2)
Outcomes	Targeted behaviour: prescribing rates for bp and reflux Baseline performance: unclear
Notes	Format: verbal and written Source: investigators Frequency: once only Instructions: no explicit target or action plan Nature of change: mix of increase and decrease or change behaviours
Risk of bias	

Nilsson 2001 *(Continued)*

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Unable to assess
Allocation concealment (selection bias)	Low risk	Cluster trial, allocation after recruitment completed
Blinding (performance bias and detection bias) All outcomes	Low risk	Prescription data used
Incomplete outcome data (attrition bias) All outcomes	Low risk	40 GPs provided data
Selective reporting (reporting bias)	Low risk	Appropriate outcomes measured
Baseline similar?	Unclear risk	Unable to assess
No contamination?	Low risk	Separate GPs
Risk of bias overall?	Unclear risk	Unable to assess

Norton 1985

Methods	Design: cluster RCT
Participants	Country: Canada Setting: primary care or outpatient Specialty: GP/family physicians N health professionals: 6 N patients: -
Interventions	Description of Groups: AF vs control (with active control)
Outcomes	Targeted behaviour: compliance with standards for GU diseases Baseline performance: unclear
Notes	Format: unclear Source: unclear Frequency: unclear Instructions: no explicit target or action plan Nature of change: mix of increase and decrease or change behaviours
Risk of bias	

Norton 1985 *(Continued)*

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Unable to assess
Allocation concealment (selection bias)	Low risk	Cluster trial, allocation after recruitment completed
Blinding (performance bias and detection bias) All outcomes	Low risk	Auditors were blinded
Incomplete outcome data (attrition bias) All outcomes	Low risk	All 6 audits completed
Selective reporting (reporting bias)	Low risk	Appropriate outcomes measured
Baseline similar?	Low risk	Table 1
No contamination?	Low risk	Separate GPs
Risk of bias overall?	Low risk	

O'Connell 1999

Methods	Design: cluster RCT
Participants	Country: Australia Setting: primary care or outpatient Specialty: GP/family physicians N health professionals: 2440 N patients: -
Interventions	Description of Groups: AF vs control
Outcomes	Targeted behaviour: prescribing rates Baseline performance: unclear
Notes	Format: written Source: rep from employer or quality assurance org Frequency: less than monthly, more than once Instructions: no explicit target or action plan Nature of change: mix of increase and decrease or change behaviours
Risk of bias	

O'Connell 1999 *(Continued)*

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Stratification with block size of 4
Allocation concealment (selection bias)	Low risk	Cluster trial, allocation after recruitment completed
Blinding (performance bias and detection bias) All outcomes	Low risk	Prescribing rates objective data
Incomplete outcome data (attrition bias) All outcomes	Low risk	Data from all subjects
Selective reporting (reporting bias)	Low risk	Appropriate outcomes measured
Baseline similar?	Low risk	Table
No contamination?	Low risk	Avoided by postal codes
Risk of bias overall?	Low risk	

O'Connor 2009

Methods	Design: cluster RCT
Participants	Country: USA Setting: primary care or outpatient Specialty: GP/family physicians, internists (and diabetes educators) N health professionals: 123 N patients: 3703
Interventions	Description of Groups: usual care vs feedback only vs feedback plus patient letters vs both
Outcomes	Targeted behaviour: diabetes management Baseline performance: unclear
Notes	Format: written Source: investigators Frequency: less than monthly, more than once Instructions: action plan provided, but no specific target Nature of change: mix of increase and decrease or change behaviours
Risk of bias	

O'Connor 2009 *(Continued)*

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Unable to assess
Allocation concealment (selection bias)	Low risk	Cluster trial, allocation after recruitment completed
Blinding (performance bias and detection bias) All outcomes	Unclear risk	Unable to assess
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Unable to assess
Selective reporting (reporting bias)	Low risk	Appropriate outcomes measured
Baseline similar?	Low risk	Table 1
No contamination?	Low risk	Separate GPs
Risk of bias overall?	Unclear risk	Unable to assess

Ornstein 2004

Methods	Design: cluster RCT
Participants	Country: USA Setting: primary care or outpatient Specialty: GP/family physicians, internists, (and mid-level providers) N health professionals: 61 N patients: 87291
Interventions	Description of Groups: feedback only vs feedback plus intensive academic detailing
Outcomes	Targeted behaviour: CVD guideline adherence Baseline performance: low
Notes	Format: written Source: investigators Frequency: less than monthly, more than once Instructions: no explicit target or action plan Nature of change: mix of increase and decrease or change behaviours
Risk of bias	

Ornstein 2004 *(Continued)*

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Baseline adaptive randomization scheme
Allocation concealment (selection bias)	Low risk	Cluster trial, allocation after recruitment completed
Blinding (performance bias and detection bias) All outcomes	Low risk	Data from computerized source
Incomplete outcome data (attrition bias) All outcomes	Low risk	3/20 subjects dropped out
Selective reporting (reporting bias)	Low risk	Appropriate outcomes measured
Baseline similar?	Unclear risk	Unable to assess
No contamination?	Low risk	Separate GPs
Risk of bias overall?	Low risk	

Palmer 1985

Methods	Design: cluster RCT
Participants	Country: USA Setting: inpatient Specialty: internists, pediatricians (and nurses) N health professionals: 711 N patients: -
Interventions	Description of Groups: feedback + education vs control
Outcomes	Targeted behaviour: variation from guideline/standard of care for 8 conditions Baseline performance: unclear
Notes	Format: verbal Source: unclear Frequency: once only Instructions: action plan provided, but no specific target Nature of change: mix of increase and decrease or change behaviours
Risk of bias	

Palmer 1985 *(Continued)*

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Blocked randomization
Allocation concealment (selection bias)	Low risk	Cluster trial, allocation after recruitment completed
Blinding (performance bias and detection bias) All outcomes	Unclear risk	Unable to assess
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	High turnover of professionals, impact unclear
Selective reporting (reporting bias)	Low risk	Appropriate outcomes measured
Baseline similar?	Unclear risk	Unable to assess
No contamination?	Low risk	Separate clusters
Risk of bias overall?	Unclear risk	Unable to assess

Phillips 2005

Methods	Design: cluster RCT
Participants	Country: USA Setting: outpatient Specialty: internists N health professionals: 345 N patients: 4138
Interventions	Description of Groups: usual care vs AF vs reminders vs both
Outcomes	Targeted behaviour: diabetes management Baseline performance: unclear
Notes	Format: verbal Source: supervisor or senior colleague Frequency: monthly Instructions: goal or target, but no action plan Nature of change: mix / unclear
Risk of bias	

Phillips 2005 *(Continued)*

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Unable to assess
Allocation concealment (selection bias)	Unclear risk	Unable to assess
Blinding (performance bias and detection bias) All outcomes	Unclear risk	Blood pressure measures may be subject to bias due to lack of blinding
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Unable to assess
Selective reporting (reporting bias)	Unclear risk	Unable to assess
Baseline similar?	Low risk	Table 1
No contamination?	Unclear risk	Unable to assess
Risk of bias overall?	Unclear risk	Unable to assess

Pimlott 2003

Methods	Design: patient or provider level RCT
Participants	Country: Canada Setting: primary care or outpatient Specialty: GP/family physicians N health professionals: 374 N patients: -
Interventions	Description of Groups: AF vs control
Outcomes	Targeted behaviour: % long acting/total benzodiazepine prescriptions Baseline performance: unclear
Notes	Format: written Source: investigators Frequency: less than monthly, more than once Instructions: explicit, measurable target but no action plan Nature of change: decrease
Risk of bias	

Pimlott 2003 *(Continued)*

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Unable to assess
Allocation concealment (selection bias)	Low risk	Cluster trial, allocation after recruitment completed
Blinding (performance bias and detection bias) All outcomes	Low risk	Data collected from central database
Incomplete outcome data (attrition bias) All outcomes	Low risk	All physicians followed up
Selective reporting (reporting bias)	Low risk	Appropriate outcomes measured
Baseline similar?	Low risk	Table 1
No contamination?	Low risk	Separate GPs
Risk of bias overall?	Low risk	

Quinley 2004

Methods	Design: cluster RCT
Participants	Country: USA Setting: primary care or outpatient Specialty: GP/family physicians, internists (and office managers) N health professionals: 950 practices N patients: -
Interventions	Description of Groups: AF vs AF + telephone follow-up
Outcomes	Targeted behaviour: immunisation rate Baseline performance: unclear
Notes	Format: verbal (verbal and written) Source: investigators Frequency: once only Instructions: action plan provided, but no specific target Nature of change: increase
Risk of bias	

Quinley 2004 *(Continued)*

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Unable to assess
Allocation concealment (selection bias)	Low risk	Cluster trial, allocation after recruitment completed
Blinding (performance bias and detection bias) All outcomes	Low risk	Data collected from central claims
Incomplete outcome data (attrition bias) All outcomes	Low risk	Data collected from Medicare, very small number had insufficient data
Selective reporting (reporting bias)	Low risk	Appropriate outcomes measured
Baseline similar?	Low risk	See page 108
No contamination?	Low risk	Randomization by practice
Risk of bias overall?	Low risk	

Raasch 2000

Methods	Design: patient or provider level RCT
Participants	Country: Australia Setting: primary care or outpatient Specialty: GP/family physicians N health professionals: 46 N patients: 1366
Interventions	Description of Groups: AF vs control
Outcomes	Targeted behaviour: % correct clinical diagnosis for skin cancer Baseline performance: moderate
Notes	Format: verbal and written Source: investigators Frequency: once only Instructions: action plan provided, but no specific target Nature of change: mix of increase and decrease or change behaviours
Risk of bias	

Raasch 2000 *(Continued)*

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Random number table
Allocation concealment (selection bias)	Low risk	Cluster trial, allocation after recruitment completed
Blinding (performance bias and detection bias) All outcomes	Unclear risk	Unable to assess
Incomplete outcome data (attrition bias) All outcomes	Low risk	4/46 lost to follow-up
Selective reporting (reporting bias)	Low risk	Appropriate outcomes measured
Baseline similar?	Unclear risk	Unable to assess
No contamination?	Unclear risk	Unable to assess
Risk of bias overall?	Unclear risk	Unable to assess

Rantz 2001

Methods	Design: cluster RCT
Participants	Country: USA Setting: primary care or outpatient Specialty: nurses N health professionals: 87 hospitals N patients: -
Interventions	Description of Groups: usual care vs AF + education vs AF + educational outreach
Outcomes	Targeted behaviour: 13 quality indicators in nursing homes Baseline performance: unclear
Notes	Format: verbal and written Source: investigators Frequency: less than monthly, more than once Instructions: explicit, measurable target and action plan Nature of change: mix of increase and decrease or change behaviours
Risk of bias	

Rantz 2001 *(Continued)*

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Unable to assess
Allocation concealment (selection bias)	Low risk	Cluster trial, allocation after recruitment completed
Blinding (performance bias and detection bias) All outcomes	Unclear risk	Unable to assess
Incomplete outcome data (attrition bias) All outcomes	Low risk	Data analyzed as intention to treat.
Selective reporting (reporting bias)	Unclear risk	Unable to assess
Baseline similar?	Low risk	See Table 1
No contamination?	Low risk	Separate nursing homes
Risk of bias overall?	Unclear risk	Unable to assess

Rask 2001

Methods	Design: cluster RCT
Participants	Country: USA Setting: primary care or outpatient Specialty: GP/family physicians, internists N health professionals: 28 (4 practices) N patients: 491
Interventions	Descriptions of Groups: AF vs AF + educational outreach by nurse plus patient reminders
Outcomes	Targeted behaviour: 6 process and 3 patient outcomes for diabetes Baseline performance: moderate
Notes	Format: written Source: rep from employer or quality assurance org Frequency: less than monthly Instructions: no explicit target/goal or action plan Nature of change: mix / unclear
Risk of bias	

Rask 2001 *(Continued)*

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Unable to assess
Allocation concealment (selection bias)	Low risk	Cluster trial, allocation after recruitment completed
Blinding (performance bias and detection bias) All outcomes	Unclear risk	Unable to assess
Incomplete outcome data (attrition bias) All outcomes	Low risk	Data on all 4 clinics
Selective reporting (reporting bias)	Unclear risk	Unable to assess
Baseline similar?	Unclear risk	Unable to assess
No contamination?	Unclear risk	Unable to assess
Risk of bias overall?	Unclear risk	Unable to assess

Robling 2002

Methods	Design: cluster RCT
Participants	Country: UK Setting: primary care or outpatient Specialty: GP/family physicians N health professionals: 37 practices N patients: -
Interventions	Description of Groups: usual care vs education vs feedback vs both
Outcomes	Targeted behaviour: % compliance with guidelines for lumbar spine and knee MRI Baseline performance: unclear
Notes	Format: written Source: investigators Frequency: once only Instructions: no explicit target or action plan Nature of change: mix of increase and decrease or change behaviours
Risk of bias	

Robling 2002 *(Continued)*

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Random number table
Allocation concealment (selection bias)	Low risk	Cluster trial, allocation after recruitment completed
Blinding (performance bias and detection bias) All outcomes	Low risk	Quote "Panel members blinded to randomization"
Incomplete outcome data (attrition bias) All outcomes	High risk	High losses
Selective reporting (reporting bias)	Low risk	Appropriate outcomes measured
Baseline similar?	Unclear risk	Unable to assess
No contamination?	Low risk	Separate GPS
Risk of bias overall?	High risk	Incomplete follow up and no baseline data

Ruangkanchanasetr 1993

Methods	Design: cluster RCT
Participants	Country: Thailand Setting: primary care or outpatient Specialty: pediatricians N health professionals: 18 N patients: -
Interventions	Description of Groups: AF vs control
Outcomes	Targeted behaviour: appropriateness of lab tests Baseline performance: unclear
Notes	Format: unclear Source: unclear Frequency: more than monthly Instructions: no explicit target or action plan Nature of change: decrease
Risk of bias	

Ruangkanchanasetr 1993 *(Continued)*

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Unable to assess
Allocation concealment (selection bias)	Low risk	Cluster trial, allocation after recruitment completed
Blinding (performance bias and detection bias) All outcomes	Unclear risk	Unable to assess
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Unable to assess, no information about followup provided
Selective reporting (reporting bias)	Low risk	Appropriate outcomes measured
Baseline similar?	Unclear risk	Unable to assess
No contamination?	Unclear risk	Unable to assess
Risk of bias overall?	Unclear risk	Unable to assess

Rubin 2001

Methods	Design: cluster RCT
Participants	Country: Australia Setting: inpatient Specialty: entire hospital involved N health professionals: unclear (10 hospitals) N patients: 1117
Interventions	Description of Groups: AF written to CEO of hospital vs AF written to CEO and presented verbally to staff
Outcomes	Targeted behaviour: appropriateness of transfusions Baseline performance: moderate
Notes	Format: both Source: investigators Frequency: once only Instructions: action plan Nature of change: decrease
Risk of bias	

Rubin 2001 *(Continued)*

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Random number table
Allocation concealment (selection bias)	Low risk	Cluster trial, allocation after recruitment completed
Blinding (performance bias and detection bias) All outcomes	Low risk	Outcome assessors blinded to intervention
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Unable to assess
Selective reporting (reporting bias)	Unclear risk	Unable to assess
Baseline similar?	Unclear risk	Unable to assess
No contamination?	Low risk	Separate hospitals
Risk of bias overall?	Unclear risk	Unable to assess

Rust 1999

Methods	Design: cluster RCT
Participants	Country: USA Setting: primary care or outpatient Specialty: pediatricians N health professionals: 32 N patients: -
Interventions	Description of Groups: AF vs control
Outcomes	Targeted behaviour: rates of immunisation Baseline performance: unclear
Notes	Format: written Source: unclear Frequency: monthly Instructions: action plan provided, but no specific target Nature of change: mix of increase and decrease or change behaviours
Risk of bias	

Rust 1999 *(Continued)*

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Unable to assess
Allocation concealment (selection bias)	Unclear risk	Unable to assess
Blinding (performance bias and detection bias) All outcomes	High risk	Not done
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Unable to assess
Selective reporting (reporting bias)	Low risk	Cluster trial, allocation after recruitment completed
Baseline similar?	Unclear risk	Unable to assess
No contamination?	Unclear risk	Unable to assess
Risk of bias overall?	High risk	

Sandbaek 1999

Methods	Design: cluster RCT
Participants	Country: Denmark Setting: primary care or outpatient Specialty: GP/family physicians N health professionals: 133 N patients: -
Interventions	Description of Groups: A F + education vs control
Outcomes	Targeted behaviour: % advised about AIDS Baseline performance: unclear
Notes	Format: written Source: unclear Frequency: once only Instructions: action plan provided, but no specific target Nature of change: mix of increase and decrease or change behaviours
Risk of bias	

Sandbaek 1999 *(Continued)*

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	"prospective randomized controlled design"
Allocation concealment (selection bias)	Unclear risk	Unable to assess
Blinding (performance bias and detection bias) All outcomes	High risk	Self report
Incomplete outcome data (attrition bias) All outcomes	Low risk	Less than 10% dropout
Selective reporting (reporting bias)	Low risk	Appropriate outcomes measure
Baseline similar?	Unclear risk	Unable to assess
No contamination?	Unclear risk	Unable to assess
Risk of bias overall?	High risk	Blinding - self-report

Sauaia 2000

Methods	Design: cluster RCT
Participants	Country: USA Setting: inpatient Specialty: internists N health professionals: 18 hospitals N patients: -
Interventions	Description of Groups: AF to just admin vs AF + opinion leader to all cardiac staff
Outcomes	Targeted behaviour: quality indicators for AMI Baseline performance: unclear
Notes	Format: verbal (verbal+written) Source: investigators Frequency: once only Instructions: no explicit target or action plan Nature of change: increase
Risk of bias	

Sauaia 2000 *(Continued)*

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Computer randomization
Allocation concealment (selection bias)	Low risk	Cluster trial, allocation after recruitment completed
Blinding (performance bias and detection bias) All outcomes	Low risk	Quality indicators collected centrally
Incomplete outcome data (attrition bias) All outcomes	Low risk	18 hospitals included in data collection
Selective reporting (reporting bias)	Low risk	Appropriate outcomes reported
Baseline similar?	Low risk	Table 2
No contamination?	Low risk	Separate hospitals
Risk of bias overall?	Low risk	

Schectman 1995

Methods	Design: patient or provider level RCT
Participants	Country: USA Setting: primary care or outpatient Specialty: GP/family physicians, internists N health professionals: 63 N patients: -
Interventions	Description of Groups: AF vs control
Outcomes	Targeted behaviour: appropriate prescribing of H2 blockers Baseline performance: low
Notes	Format: written Source: rep from employer or quality assurance org Frequency: once only Instructions: no explicit target or action plan Nature of change: mix of increase and decrease or change behaviours
Risk of bias	

Schectman 1995 *(Continued)*

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Unable to assess
Allocation concealment (selection bias)	Low risk	Cluster trial, allocation after recruitment completed
Blinding (performance bias and detection bias) All outcomes	Unclear risk	Unable to assess
Incomplete outcome data (attrition bias) All outcomes	Low risk	All physicians included in analysis
Selective reporting (reporting bias)	Low risk	Appropriate outcomes reported
Baseline similar?	Low risk	See Figure 1
No contamination?	Unclear risk	Unable to assess
Risk of bias overall?	Unclear risk	Unable to assess

Schectman 2003

Methods	Design: cluster RCT
Participants	Country: USA Setting: primary care or outpatient Specialty: GP/family physicians, internists, nurse practioners, physician assistants N health professionals: 85 N patients: -
Interventions	Description of Groups: usual care vs patient education vs feedback + opinion leader + education vs both
Outcomes	Targeted behaviour: % compliance with guidelines for low back pain Baseline performance: unclear
Notes	Format: verbal and written Source: investigators Frequency: less than monthly, more than once Instructions: no explicit target or action plan Nature of change: mix of increase and decrease or change behaviours
Risk of bias	

Schectman 2003 *(Continued)*

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Sealed envelopes used
Allocation concealment (selection bias)	Low risk	Cluster trial, allocation after recruitment completed
Blinding (performance bias and detection bias) All outcomes	Low risk	Abstractors blinded
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Unable to assess
Selective reporting (reporting bias)	Unclear risk	Unable to assess
Baseline similar?	Unclear risk	Unable to assess
No contamination?	Low risk	Separate sites for GPs
Risk of bias overall?	Unclear risk	Unable to assess

Schneider 2008

Methods	Design: cluster RCT
Participants	Country: Germany Setting: primary care or outpatient Specialty: GP/family physicians N health professionals: 42 practices N patients: 185
Interventions	Description of Groups: quality circles with feedback vs quality circles w feedback identifying top performers
Outcomes	Targeted behaviour: asthma symptom control Baseline performance: moderate
Notes	Format: verbal and written Source: investigators Frequency: once only Instructions: no explicit target or action plan Nature of change: mix of increase and decrease or change behaviours
Risk of bias	

Schneider 2008 *(Continued)*

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Unable to assess
Allocation concealment (selection bias)	Low risk	Cluster trial, allocation after recruitment completed
Blinding (performance bias and detection bias) All outcomes	Unclear risk	Unable to assess
Incomplete outcome data (attrition bias) All outcomes	High risk	Many lost to follow-up
Selective reporting (reporting bias)	High risk	Only secondary outcomes
Baseline similar?	Unclear risk	Unable to assess
No contamination?	Unclear risk	Unable to assess
Risk of bias overall?	High risk	Possible contamination, blinding unclear, only secondary outcomes, not good follow-up

Scholes 2006

Methods	Design: cluster RCT
Participants	Country: USA Setting: primary care or outpatient Specialty: GP/family physicians, internists, pediatrics (and nurses) N health professionals: 204 N patients: 11755
Interventions	Description of Groups: passive guideline dissemination vs active dissemination w opinion leaders, reminders and feedback
Outcomes	Targeted behaviour: chlamydia screening Baseline performance: moderate
Notes	Format: written Source: investigators Frequency: less than monthly, more than once Instructions: no explicit target or action plan Nature of change: increase

Scholes 2006 *(Continued)*

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Unable to assess
Allocation concealment (selection bias)	Low risk	Cluster trial, allocation after recruitment completed
Blinding (performance bias and detection bias) All outcomes	Low risk	Database used for data
Incomplete outcome data (attrition bias) All outcomes	Low risk	Data from all clinics presented
Selective reporting (reporting bias)	Unclear risk	Unable to assess
Baseline similar?	Low risk	Table 2
No contamination?	Low risk	Separate clinics
Risk of bias overall?	Low risk	

Sinclair 1982

Methods	Design: cluster RCT
Participants	Country: Canada Setting: primary care or outpatient Specialty: mental health clinicians, management N health professionals: 11 N patients: -
Interventions	Description of Groups: AF + education vs control
Outcomes	Targeted behaviour: quality scores for mental health services Baseline performance: moderate
Notes	Format: unclear Source: supervisor Frequency: once only Instructions: action plan provided, but no specific target Nature of change: mix of increase and decrease or change behaviours

Sinclair 1982 *(Continued)*

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Unable to assess
Allocation concealment (selection bias)	Low risk	Cluster trial, allocation after recruitment completed
Blinding (performance bias and detection bias) All outcomes	Low risk	Data collected by blinded assessor
Incomplete outcome data (attrition bias) All outcomes	Low risk	Follow up for all professionals
Selective reporting (reporting bias)	Low risk	Appropriate outcomes reported
Baseline similar?	Unclear risk	Unable to assess
No contamination?	Unclear risk	Unable to assess
Risk of bias overall?	Unclear risk	Unable to assess

Siriwardena 2002

Methods	Design: cluster RCT
Participants	Country: UK Setting: primary care or outpatient Specialty: GP/family physician, nurse, management N health professionals: 30 practices N patients: -
Interventions	Description of Groups: AF vs AF + educational outreach
Outcomes	Targeted behaviour: vaccination rates Baseline performance: unclear
Notes	Format: verbal and written (groupverbal),written (groupwritten) Source: investigators Frequency: once only Instructions: no explicit target or action plan Nature of change: increase

Siriwardena 2002 *(Continued)*

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Unable to assess
Allocation concealment (selection bias)	Low risk	Cluster trial, allocation after recruitment completed
Blinding (performance bias and detection bias) All outcomes	Low risk	Data collected as part of national campaign
Incomplete outcome data (attrition bias) All outcomes	Low risk	Data from all practices provided
Selective reporting (reporting bias)	Low risk	Appropriate outcomes reported
Baseline similar?	Low risk	Groups similar at baseline
No contamination?	Low risk	Separate practices
Risk of bias overall?	Low risk	

Smith 1998

Methods	Design: cluster RCT
Participants	Country: USA Setting: unclear Specialty: unclear N health professionals: 222 N patients: -
Interventions	Description of Groups: AF vs control
Outcomes	Targeted behaviour: prescribing of benzodiazepines Baseline performance: unclear
Notes	Format: written Source: rep from employer or quality assurance org Frequency: once only Instructions: no explicit target or action plan Nature of change: decrease

Smith 1998 *(Continued)*

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Computer randomization
Allocation concealment (selection bias)	Low risk	Cluster trial, allocation after recruitment completed
Blinding (performance bias and detection bias) All outcomes	Low risk	Data collected through Medicaid
Incomplete outcome data (attrition bias) All outcomes	Low risk	Less than 10% stopped taking drugs
Selective reporting (reporting bias)	Low risk	Appropriate outcomes reported
Baseline similar?	Low risk	Table 1
No contamination?	Low risk	Randomized by cluster
Risk of bias overall?	Low risk	

Socolar 1998

Methods	Design: patient or provider level RCT
Participants	Country: USA Setting: primary care or outpatient Specialty: pediatricians, psychiatrists, psychologists N health professionals: 147 N patients: -
Interventions	Description of Groups: AF vs control
Outcomes	Targeted behaviour: documentation in medical records for sexual abuse Baseline performance: unclear
Notes	Format: written Source: investigators Frequency: once only Instructions: action plan provided, but no specific target Nature of change: mix of increase and decrease or change behaviours

Socolar 1998 *(Continued)*

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Random number assignment
Allocation concealment (selection bias)	Low risk	Cluster trial, allocation after recruitment completed.
Blinding (performance bias and detection bias) All outcomes	Low risk	Documentation assessed by blinded reviewers
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Unable to assess
Selective reporting (reporting bias)	Low risk	Appropriate outcomes reported
Baseline similar?	Low risk	Table 1
No contamination?	Low risk	Separate professionals
Risk of bias overall?	Unclear risk	Unable to assess

Solomon 2004

Methods	Design: cluster RCT
Participants	Country: USA Setting: primary care or outpatient Specialty: internists N health professionals: 21 N patients: 373
Interventions	Description of Groups: usual care vs education with feedback
Outcomes	Targeted behaviour: osteoporosis management Baseline performance: low
Notes	Format: verbal and written Source: investigators Frequency: once only Instructions: no explicit target or action plan Nature of change: increase

Solomon 2004 *(Continued)*

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Unable to assess
Allocation concealment (selection bias)	Low risk	Cluster trial, allocation after recruitment completed
Blinding (performance bias and detection bias) All outcomes	Unclear risk	Unable to assess
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Unable to assess
Selective reporting (reporting bias)	Low risk	Appropriate outcomes reported
Baseline similar?	Low risk	One physician lost to followup, not related to intervention. See Table 1
No contamination?	Unclear risk	Unable to assess
Risk of bias overall?	Unclear risk	Unable to assess

Sommers 1984

Methods	Design: cluster RCT
Participants	Country: USA Setting: inpatient Specialty: internists, surgeons N health professionals: 103 N patients: -
Interventions	Description of Groups: usual care vs feedback vs feedback with consensus process
Outcomes	Targeted behaviour: compliance with guidelines for anaemia Baseline performance: moderate
Notes	Format: verbal Source: unclear Frequency: once only Instructions: no explicit target or action plan Nature of change: mix of increase and decrease or change behaviours

Sommers 1984 *(Continued)*

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Unable to assess
Allocation concealment (selection bias)	Low risk	Cluster trial, allocation after recruitment completed
Blinding (performance bias and detection bias) All outcomes	Unclear risk	Unable to assess
Incomplete outcome data (attrition bias) All outcomes	High risk	Data from one hospital not included
Selective reporting (reporting bias)	Low risk	Appropriate outcomes reported
Baseline similar?	Unclear risk	Unable to assess
No contamination?	Unclear risk	Unable to assess
Risk of bias overall?	High risk	Difficulty with follow-up data

Soumerai 1998

Methods	Design: cluster RCT
Participants	Country: USA Setting: inpatient Specialty: surgeons (and entire hospital) N health professionals: 37 hospitals N patients: 2938
Interventions	Description of Groups: AF vs AF + opinion leader + education
Outcomes	Targeted behaviour: appropriate prescribing post MI Baseline performance: uncllear
Notes	Format: written Source: investigators Frequency: once only Instructions: no explicit target or action plan Nature of change: mix of increase and decrease or change behaviours

Soumerai 1998 *(Continued)*

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Unable to assess
Allocation concealment (selection bias)	Low risk	Cluster trial, allocation after recruitment completed
Blinding (performance bias and detection bias) All outcomes	Unclear risk	Unable to assess
Incomplete outcome data (attrition bias) All outcomes	Low risk	See Figure 1
Selective reporting (reporting bias)	Low risk	Appropriate outcomes reported
Baseline similar?	Low risk	See Table 1
No contamination?	Low risk	Hospitals are separate units
Risk of bias overall?	Unclear risk	Unable to assess

Svetkey 2009

Methods	Design: cluster RCT
Participants	Country: USA Setting: primary care or outpatient Specialty: GP/family physicians, internists N health professionals: 32 N patients: 574
Interventions	Description of Groups: usual care vs education and quarterly feedback vs group visits and pt self mgmt vs both
Outcomes	Targeted behaviour: htn management Baseline performance: moderate
Notes	Format: written Source: investigators Frequency: less than monthly, more than once Instructions: no explicit target or action plan Nature of change: mix of increase and decrease or change behaviours

Svetkey 2009 *(Continued)*

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Matched pair randomization
Allocation concealment (selection bias)	Low risk	Cluster trial, allocation after recruitment completed
Blinding (performance bias and detection bias) All outcomes	Unclear risk	Unable to assess
Incomplete outcome data (attrition bias) All outcomes	Low risk	Less than 15% dropout, comparable in all groups
Selective reporting (reporting bias)	Low risk	Appropriate outcomes reported
Baseline similar?	Low risk	See Table 1
No contamination?	Low risk	GPs separate
Risk of bias overall?	Unclear risk	Unable to assess

Søndergaard 2002

Methods	Design: cluster RCT
Participants	Country: Denmark Setting: primary care or outpatient Specialty: GP/family physicians N health professionals: 292 N patients: -
Interventions	Description of Groups: usual care vs AF (aggregated data) vs AF (with individual pt data)
Outcomes	Targeted behaviour: % treated with inhaled steroids Baseline performance: unclear
Notes	Format: written Source: investigators Frequency: less than monthly, more than once Instructions: no explicit target or action plan Nature of change: mix of increase and decrease or change behaviours

Søndergaard 2002 *(Continued)*

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Unable to assess
Allocation concealment (selection bias)	Low risk	Cluster trial, allocation after recruitment completed
Blinding (performance bias and detection bias) All outcomes	Low risk	Data collected from billing database
Incomplete outcome data (attrition bias) All outcomes	Low risk	All prescriptions assessed
Selective reporting (reporting bias)	Low risk	Appropriate outcomes reported
Baseline similar?	Low risk	Quote "There was no statistically significant differences between the intervention and the control groups....at the onset of the trial."
No contamination?	Low risk	Practices randomized
Risk of bias overall?	Low risk	

Søndergaard 2003

Methods	Design: cluster RCT
Participants	Country: Denmark Setting: primary care or outpatient Specialty: GP/family physicians N health professionals: 299 N patients: 455843
Interventions	Description of Groups: AF vs control
Outcomes	Targeted behaviour: % prescribtions for narrow-spectrum penicillins Baseline performance: unclear
Notes	Format: written Source: investigators Frequency: once only Instructions: no explicit target or action plan Nature of change: decrease

Søndergaard 2003 *(Continued)*

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Unable to assess
Allocation concealment (selection bias)	Low risk	Cluster trial, allocation after recruitment completed
Blinding (performance bias and detection bias) All outcomes	Low risk	Data collected from billing database
Incomplete outcome data (attrition bias) All outcomes	Low risk	All prescriptions assessed
Selective reporting (reporting bias)	Low risk	Appropriate outcomes reported
Baseline similar?	Low risk	Table 1
No contamination?	Low risk	Practices randomized
Risk of bias overall?	Low risk	

Søndergaard 2006

Methods	Design: cluster RCT
Participants	Country: Denmark Setting: primary care or outpatient Specialty: GP/family physicians N health professionals: 28 N patients: 320
Interventions	Description of Groups: usual care vs academic detailing featuring feedback
Outcomes	Targeted behaviour: prescribing rates for heart disease Baseline performance: moderate
Notes	Format: verbal and written Source: investigators Frequency: less than monthly, more than once Instructions: no explicit target or action plan Nature of change: increase

Søndergaard 2006 *(Continued)*

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Random number table
Allocation concealment (selection bias)	Low risk	Cluster trial, allocation after recruitment completed
Blinding (performance bias and detection bias) All outcomes	High risk	Audit done by participants
Incomplete outcome data (attrition bias) All outcomes	Low risk	Equal dropouts in both groups, see Figure 1
Selective reporting (reporting bias)	Low risk	Appropriate outcomes reported
Baseline similar?	Low risk	See Table 1
No contamination?	Unclear risk	Unable to assess
Risk of bias overall?	High risk	Outcomes assessed by participants, unclear blinding

Thomas 2006

Methods	Design: cluster RCT
Participants	Country: UK Setting: primary care or outpatient Specialty: GP/family physicians N health professionals: 370 (85 practices) N patients: unclear
Interventions	Description of Groups: usual care vs AF with educational messages vs educational reminders vs both
Outcomes	Targeted behaviour: number of laboratory tests ordered Baseline performance: unclear
Notes	Format: written Source: investigators Frequency: less than monthly Instructions: action plan Nature of change: decrease

Thomas 2006 *(Continued)*

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Quote "cluster randomization...with a minimization procedure"
Allocation concealment (selection bias)	Low risk	Cluster trial, allocation after recruitment completed
Blinding (performance bias and detection bias) All outcomes	Low risk	Quote "The laboratory personnel who processed the requests were unaware of the intervention-group status."
Incomplete outcome data (attrition bias) All outcomes	Low risk	Data from all sites randomized included in analysis
Selective reporting (reporting bias)	Low risk	Appropriate outcomes reported
Baseline similar?	Unclear risk	Unable to assess
No contamination?	Low risk	Separate practices
Risk of bias overall?	Low risk	

Thomas 2007

Methods	Design: cluster RCT
Participants	Country: USA Setting: primary care or outpatient Specialty: internists N health professionals: 78 N patients: 483
Interventions	Description of Groups: usual care vs AF with education and reminders
Outcomes	Targeted behaviour: management of diabetes Baseline performance: unclear
Notes	Format: written Source: investigators Frequency: less than monthly Instructions: no explicit target or action plan Nature of change: mixed or unclear

Thomas 2007 *(Continued)*

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Unable to assess
Allocation concealment (selection bias)	Unclear risk	Unable to assess
Blinding (performance bias and detection bias) All outcomes	Unclear risk	Unclear if blood pressure assessors were blinded
Incomplete outcome data (attrition bias) All outcomes	Low risk	All subjects included in analysis
Selective reporting (reporting bias)	Low risk	Appropriate outcomes reported
Baseline similar?	Low risk	Quote "Resident demographic data (age, sex, and year in training) were similar between groups."
No contamination?	Unclear risk	Unable to assess
Risk of bias overall?	Unclear risk	Unable to assess

Tierney 1986

Methods	Design: cluster RCT
Participants	Country: USA Setting: primary care or outpatient Specialty: internists N health professionals: 135 N patients: -
Interventions	Description of Groups: usual care vs feedback vs reminders vs both
Outcomes	Targeted behaviour: % pts who received preventive care according to guidelines Baseline performance: low
Notes	Format: written Source: unclear Frequency: monthly Instructions: action plan provided, but no specific target Nature of change: increase

Tierney 1986 *(Continued)*

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Unable to assess
Allocation concealment (selection bias)	Unclear risk	Unable to assess
Blinding (performance bias and detection bias) All outcomes	Unclear risk	Unable to assess
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Unable to assess
Selective reporting (reporting bias)	Low risk	Appropriate outcomes reported
Baseline similar?	Unclear risk	No baseline data presented
No contamination?	Unclear risk	Unable to assess
Risk of bias overall?	Unclear risk	Unable to assess

Tu 2009

Methods	Design: cluster RCT
Participants	Country: Canada Setting: inpatient Specialty: internists (and entire hospital) N health professionals: 81 hospitals N patients: 15997
Interventions	Description of Groups: delayed feedback (usual care) vs public release of feedback (report cards)
Outcomes	Targeted behaviour: compliance w process of care indicators for CHF and MI Baseline performance: moderate
Notes	Format: written Source: investigators Frequency: once only Instructions: no explicit target or action plan Nature of change: mix of increase and decrease or change behaviours

Tu 2009 *(Continued)*

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Stratified and performed by a statistitian
Allocation concealment (selection bias)	Low risk	Cluster trial, allocation after recruitment completed
Blinding (performance bias and detection bias) All outcomes	Low risk	Blinded based on communication with author
Incomplete outcome data (attrition bias) All outcomes	Low risk	See Figure
Selective reporting (reporting bias)	Low risk	Appropriate outcomes reported
Baseline similar?	Low risk	See Table 1
No contamination?	Low risk	Hospitals separate
Risk of bias overall?	Low risk	

Van den Hombergh 1999

Methods	Design: cluster RCT
Participants	Country: Netherlands Setting: primary care or outpatient Specialty: GP/family physicians N health professionals: 90 N patients: -
Interventions	Description of Groups: AF by peer vs AF by non-physician
Outcomes	Targeted behaviour: 208 indicators of practice management Baseline performance: unclear
Notes	Format: verbal and written Source: unclear Frequency: once only Instructions: action plan provided, but no specific target Nature of change: increase

Van den Hombergh 1999 *(Continued)*

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Unable to assess
Allocation concealment (selection bias)	Low risk	Cluster trial, allocation after recruitment completed
Blinding (performance bias and detection bias) All outcomes	Unclear risk	Unable to assess
Incomplete outcome data (attrition bias) All outcomes	Low risk	All practices provided data
Selective reporting (reporting bias)	Low risk	Appropriate outcomes reported
Baseline similar?	Unclear risk	Unable to assess
No contamination?	Low risk	Practices separate
Risk of bias overall?	Unclear risk	Unable to assess

Van der Weijden 1999

Methods	Design: cluster RCT
Participants	Country: Netherlands Setting: primary care or outpatient Specialty: GP/family physicians N health professionals: 32 N patients: -
Interventions	Description of Groups: feedback + educational outreach + opinion leaders vs control
Outcomes	Targeted behaviour: compliance with guidelines for cholesterol Baseline performance: unclear
Notes	Format: verbal Source: investigators Frequency: less than monthly, more than once Instructions: action plan provided, but no specific target Nature of change: mix of increase and decrease or change behaviours

Van der Weijden 1999 *(Continued)*

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Blocked randomization
Allocation concealment (selection bias)	Low risk	Cluster trial, allocation after recruitment completed
Blinding (performance bias and detection bias) All outcomes	Low risk	Blinded data collectors
Incomplete outcome data (attrition bias) All outcomes	Low risk	20/20 practices
Selective reporting (reporting bias)	Low risk	Appropriate outcomes reported
Baseline similar?	Low risk	See Table 2
No contamination?	Low risk	Practices separate
Risk of bias overall?	Low risk	

Veninga 1999

Methods	Design: cluster RCT
Participants	Country: Netherlands, Sweden, Norway, Slovakia Setting: primary care or outpatient Specialty: GP/family physicians N health professionals: 565 N patients: -
Interventions	Description of Groups: AF with small group education vs control
Outcomes	Targeted behaviour: prescribing practices for asthma Baseline performance: unclear
Notes	Format: verbal and written Source: unclear Frequency: once only Instructions: no explicit target or action plan Nature of change: mix of increase and decrease or change behaviours

Veninga 1999 *(Continued)*

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Unable to assess
Allocation concealment (selection bias)	Unclear risk	Unable to assess
Blinding (performance bias and detection bias) All outcomes	Low risk	Quote: "Prescription data of a 6-mo (NL, S, SK) or 12-mo period (N) were collected before and after the intervention"through pharmacies, insurance companies, or directly from computerized databases of doctors dispensing drugs in their practice."
Incomplete outcome data (attrition bias) All outcomes	Low risk	Pharmacy data used
Selective reporting (reporting bias)	Low risk	Appropriate outcomes reported
Baseline similar?	Low risk	See Table 1
No contamination?	Low risk	Independent sites
Risk of bias overall?	Unclear risk	Unable to assess

Verstappen 2003

Methods	Design: cluster RCT
Participants	Country: Netherlands Setting: primary care or outpatient Specialty: GP/family physicians N health professionals: 174 (26 practices) N patients: -
Interventions	Description of Groups: AF with small group education vs control
Outcomes	Targeted behaviour: % with guidelines for tests ordering Baseline performance: unclear
Notes	Format: written Source: investigators Frequency: less than monthly, more than once Instructions: explicit, measurable target and action plan Nature of change: decrease

Verstappen 2003 *(Continued)*

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Blocked randomization
Allocation concealment (selection bias)	Low risk	Cluster trial, allocation after recruitment completed
Blinding (performance bias and detection bias) All outcomes	Low risk	Tests ordered objective data
Incomplete outcome data (attrition bias) All outcomes	Low risk	1/26 practices lost to follow-up
Selective reporting (reporting bias)	Unclear risk	Unable to assess
Baseline similar?	Low risk	See Table 2
No contamination?	Low risk	Independent clinics
Risk of bias overall?	Low risk	

Verstappen 2004

Methods	Design: cluster RCT
Participants	Country: Netherlands Setting: primary care or outpatient Specialty: GP/family physicians N health professionals: 174 (27 practices) N patients: -
Interventions	Description of Groups: AF vs AF plus learning collaboratives with cqi
Outcomes	Targeted behaviour: % with guidelines for tests ordering Baseline performance: unclear
Notes	Format: written Source: investigators Frequency: less than monthly, more than once Instructions: explicit, measurable target and action plan Nature of change: decrease

Verstappen 2004 *(Continued)*

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Blocked randomization
Allocation concealment (selection bias)	Low risk	Cluster trial, allocation after recruitment completed
Blinding (performance bias and detection bias) All outcomes	Low risk	Tests ordered objective data
Incomplete outcome data (attrition bias) All outcomes	Low risk	26/27 sites competed trial
Selective reporting (reporting bias)	Unclear risk	Unable to assess
Baseline similar?	Low risk	See Table 2, non significant differences
No contamination?	Low risk	Independent sites
Risk of bias overall?	Low risk	

Vingerhoets 2001

Methods	Design: cluster RCT
Participants	Country: Netherlands Setting: primary care or outpatient Specialty: GP/family physicians N health professionals: 55 N patients: 7286
Interventions	Description of Groups: AF vs control
Outcomes	Targeted behaviour: mean scores of patients satisfaction with general care Baseline performance: unclear
Notes	Format: written Source: investigators Frequency: once only Instructions: no explicit target or action plan Nature of change: mix of increase and decrease or change behaviours

Vingerhoets 2001 *(Continued)*

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Computer generated list of random numbers
Allocation concealment (selection bias)	Low risk	Cluster trial, allocation after recruitment completed
Blinding (performance bias and detection bias) All outcomes	Low risk	Quote "Patients were blinded for the intervention.."
Incomplete outcome data (attrition bias) All outcomes	Low risk	Less than 10% drop out, 4 physicians dropped out of control, one out of intervention
Selective reporting (reporting bias)	Low risk	Appropriate outcomes reported
Baseline similar?	Low risk	Patient groups did not differ between arms
No contamination?	Low risk	43 separate practices
Risk of bias overall?	Low risk	

Wadland 2007

Methods	Design: cluster RCT
Participants	Country: USA Setting: primary care or outpatient Specialty: GP/family physicians, internists, OBGYN N health professionals: 308 N patients: -
Interventions	Description of Groups: mailed quarterly reminders vs quarterly feedback reports
Outcomes	Targeted behaviour: smoking cessation referrals Baseline performance: moderate
Notes	Format: written Source: rep from employer or quality assurance org Frequency: less than monthly, more than once Instructions: explicit, measurable target but no action plan Nature of change: increase

Wadland 2007 *(Continued)*

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Unable to assess
Allocation concealment (selection bias)	Low risk	Cluster trial, allocation after recruitment completed
Blinding (performance bias and detection bias) All outcomes	Unclear risk	Unable to assess
Incomplete outcome data (attrition bias) All outcomes	Low risk	Over 300/308 clinicians provided data
Selective reporting (reporting bias)	Low risk	Appropriate outcomes reported
Baseline similar?	Low risk	See Table 1 and 2
No contamination?	Low risk	Separate clinics
Risk of bias overall?	High risk	Apparent conflict of interest

Wahlström 2003

Methods	Design: cluster RCT
Participants	Country: Lao Setting: primary care or outpatient.inpatient Specialty: internists and pediatricians N health professionals: 122 N patients: -
Interventions	Description of Groups: AF vs control
Outcomes	Targeted behaviour: mean indicator score for malaria, diarrhoea and pneumonia Baseline performance: unclear
Notes	Format: verbal Source: rep from employer or quality assurance org Frequency: monthly Instructions: action plan provided, but no specific target Nature of change: mix of increase and decrease or change behaviours

Wahlström 2003 *(Continued)*

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Unable to assess
Allocation concealment (selection bias)	Low risk	Cluster trial, allocation after recruitment completed
Blinding (performance bias and detection bias) All outcomes	Unclear risk	Unable to assess
Incomplete outcome data (attrition bias) All outcomes	Low risk	Data from all hospital departments provided
Selective reporting (reporting bias)	Low risk	Appropriate outcomes reported
Baseline similar?	Unclear risk	Unable to assess
No contamination?	Unclear risk	Unable to assess
Risk of bias overall?	Unclear risk	Unable to assess

Ward 1996

Methods	Design: cluster RCT
Participants	Country: Australia Setting: primary care or outpatient Specialty: GP/family physicians N health professionals: 139 N patients: 386
Interventions	Description of Groups: AF vs AF + educational outreach by peer vs AF + educational outreach by nurse
Outcomes	Targeted behaviour: compliance with guidelines for diabetes Baseline performance: unclear
Notes	Format: verbal and written Source: unclear Frequency: once only Instructions: explicit, measurable target but no action plan Nature of change: increase

Ward 1996 *(Continued)*

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Quote "randomly allocated...stratified by number of patients recruited"
Allocation concealment (selection bias)	Low risk	Cluster trial, allocation after recruitment completed
Blinding (performance bias and detection bias) All outcomes	Unclear risk	Unable to assess
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Unable to assess
Selective reporting (reporting bias)	Low risk	Appropriate outcomes reported
Baseline similar?	Unclear risk	Unable to assess
No contamination?	Unclear risk	Unable to assess
Risk of bias overall?	Unclear risk	Unable to assess

Weitzman 2009

Methods	Design: cluster RCT
Participants	Country: Israel Setting: outpatient Specialty: GP or family physician N health professionals: unclear, (4 practices) N patients: 429
Interventions	Description of Groups: AF vs AF plus patient reminders
Outcomes	Targeted behaviour: control of risk factors in diabetes Baseline performance: unclear
Notes	Format: verbal Source: investigators Frequency: once only Instructions: no goal-setting or action plan Nature of change: unclear or mix of increase and decrease behaviours

Weitzman 2009 *(Continued)*

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Unable to assess
Allocation concealment (selection bias)	Unclear risk	Unable to assess
Blinding (performance bias and detection bias) All outcomes	Low risk	Data collected from computerized medical record
Incomplete outcome data (attrition bias) All outcomes	Low risk	Data provided on all patients
Selective reporting (reporting bias)	Unclear risk	Unable to assess
Baseline similar?	Unclear risk	Unable to assess
No contamination?	Unclear risk	Unable to assess
Risk of bias overall?	Unclear risk	Unable to assess

Winickoff 1984

Methods	Design: cluster RCT
Participants	Country: USA Setting: primary care or outpatient Specialty: internists N health professionals: 16 N patients: -
Interventions	Description of Groups: AF vs control
Outcomes	Targeted behaviour: % FOBT performed Baseline performance: moderate
Notes	Format: written Source: unclear Frequency: monthly Instructions: explicit, measurable target but no action plan Nature of change: increase

Winickoff 1984 *(Continued)*

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Randomized after stratification
Allocation concealment (selection bias)	Low risk	Cluster trial, allocation after recruitment completed
Blinding (performance bias and detection bias) All outcomes	Unclear risk	Unable to assess
Incomplete outcome data (attrition bias) All outcomes	Low risk	
Selective reporting (reporting bias)	Low risk	Appropriate outcomes reported
Baseline similar?	Low risk	Performance similar pre-intervention
No contamination?	High risk	Authors acknowledge possibility
Risk of bias overall?	Unclear risk	Unable to assess

Winkens 1995

Methods	Design: patient or provider level RCT
Participants	Country: Netherlands Setting: primary care or outpatient Specialty: GP/family physicians N health professionals: 79 N patients: -
Interventions	Description of Groups: AF vs control
Outcomes	Targeted behaviour: mean non-rational tests per doctor Baseline performance: unclear
Notes	Format: written Source: supervisor Frequency: less than monthly, more than once Instructions: action plan provided, but no specific target Nature of change: mix of increase and decrease or change behaviours

Winkens 1995 *(Continued)*

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Unable to assess
Allocation concealment (selection bias)	Low risk	Cluster trial, allocation after recruitment completed
Blinding (performance bias and detection bias) All outcomes	High risk	No blinding done
Incomplete outcome data (attrition bias) All outcomes	Low risk	Data collected from central registry routinely
Selective reporting (reporting bias)	Low risk	Appropriate outcomes reported
Baseline similar?	Unclear risk	Unable to assess
No contamination?	Unclear risk	Unable to assess
Risk of bias overall?	High risk	Blinding

Wones 1987

Methods	Design: patient or provider level RCT
Participants	Country: USA Setting: inpatient Specialty: internists N health professionals: 21 N patients: -
Interventions	Description of Groups: usual care vs AF vs AF with peer comparison
Outcomes	Targeted behaviour: lab tests per patient-day Baseline performance: unclear
Notes	Format: written Source: unclear Frequency: monthly Instructions: no explicit target or action plan Nature of change: decrease

Wones 1987 *(Continued)*

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Unable to assess
Allocation concealment (selection bias)	Low risk	Cluster trial, allocation after recruitment completed
Blinding (performance bias and detection bias) All outcomes	Low risk	Utilization data from central computer
Incomplete outcome data (attrition bias) All outcomes	Low risk	Data provided on all 21 residents (Table 2)
Selective reporting (reporting bias)	Low risk	Appropriate outcomes reported
Baseline similar?	Unclear risk	Unable to assess
No contamination?	Unclear risk	Unable to assess
Risk of bias overall?	Unclear risk	Unable to assess

Ziemer 2006

Methods	Design: cluster RCT
Participants	Country: USA Setting: outpatient Specialty: internists N health professionals: 345 N patients: 4038
Interventions	Description of Groups: usual care vs AF vs reminders vs both
Outcomes	Targeted behaviour: diabetes visits with action taken to reduce glucose Baseline performance: moderate
Notes	Format: both verbal and written Source: supervisor or senior colleague Frequency: monthly Instructions: both goal-setting and action plan Nature of change: increase

Ziemer 2006 *(Continued)*

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Unable to assess
Allocation concealment (selection bias)	Unclear risk	Unable to assess
Blinding (performance bias and detection bias) All outcomes	Unclear risk	Unable to assess
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Unable to assess
Selective reporting (reporting bias)	Unclear risk	Unable to assess
Baseline similar?	Low risk	Patient groups similar, see page 509
No contamination?	Unclear risk	Unable to assess
Risk of bias overall?	Unclear risk	Unable to assess

Zwar 1999

Methods	Design: patient or provider level RCT
Participants	Country: Australia Setting: primary care or outpatient Specialty: GP/family physicians N health professionals: 157 N patients: -
Interventions	Description of Groups: AF + educational outreach vs control
Outcomes	Targeted behaviour: rate of antibiotic prescribing for urti Baseline performance: unclear
Notes	Format: written Source: investigators Frequency: once only Instructions: no explicit target or action plan Nature of change: decrease

Zwar 1999 *(Continued)*

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Unable to assess
Allocation concealment (selection bias)	Low risk	Cluster trial, allocation after recruitment completed
Blinding (performance bias and detection bias) All outcomes	High risk	Self report data
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Unable to assess impact of dropout
Selective reporting (reporting bias)	Low risk	Appropriate outcomes reported
Baseline similar?	Low risk	Trainees similar at baseline
No contamination?	Unclear risk	Unable to assess
Risk of bias overall?	High risk	Lack of blinding

ACE: angiotensin-converting-enzyme (inhibitor)
AF: audit and feedback
AMI: acute myocardial infarction
BB: beta blocker
CABG: coronary artery bypass graft
CHD: coronary heart disease
CHF: congestive heart failure
CQI: continuous quality improvement
CV: cardiovascular
DVT: deep vein thrombosis
EMR: electronic medical record
ER: emergency room
FOBT: faecal occult blood test
MI: myocardial infarction
MRI: magnetic resonance imaging
NSAID: non-steroidal inflammatory drug
RCT: randomised controlled trial
Rx; treatment
QA: quality assurance
QI: quality improvement
UTI: urinary tract infection
vs: versus
VTE: venous thromboembolism

Characteristics of excluded studies *[ordered by study ID]*

Study	Reason for exclusion
Aittasalo 2006	Not feedback

Study	Reason for exclusion
Allard 2001	Not feedback
Allison 2005	Not core
Althabe 2008	Not core
Anderson 1996	Not randomised
Anonymous I 1990	Not audit and feedback
Aspy 2008	Not core
Ballard 2002	Not RCT
Belcher 1990	Not core feedback
Bertoni 2009	Not core
Berwick 1986	Randomisation not specified
Billi 1987	Not audit and feedback
Bindels 2003	Not feedback
Bischoff 2000	Not RCT
Bonds 2009	Not core
Bonetti 2005	No results
Brand 2005	Not RCT
Britton 1991	Not feedback
Brown 1988	Not RCT
Buekens 1993	Not RCT
Bunting 2004	Not RCT
Campbell 2006	Not core
Carney 1992	Not feedback on performance
Chin 2007	Not core
Chowdhury 2007	Not RCT
Cleveringa 2008	Not core
Cohen 1996	Not RCT
Colón-Emeric 2007	Not core
Cope 1986	Not RCT

Study	Reason for exclusion
Cranney 1999	Not core
Crits-Christoph 2010	Not feedback
Crotty 2004	Not core
Curtis 2009	Not core
De Silva 1994	Outcome was based on self-report
Del Mar 1998	Not audit and feedback
Denton 2001	Not RCT
Dickinson 1981	Not randomised
Doherty 2006	Not core
Doherty 2007	Not a randomised trial
Dranitsaris 1995	Not feedback
Dulko 2010	Not RCT
Echouffo-Tcheugui 2009	No results
Elnicki 1998	No results
Everett	Insufficient data on results
Fallowfield 2002	'Feedback' focused on skills
Feder 1995	Not core
Ferreira 2005	Only two groups randomised
Fick 2004	Not core
Fihn 2004	Outcome not professional practice or patient outcome
Finkelstein 2001	Not core
Finkelstein 2005	Feedback not core
Finkelstein 2008	Not core
Frame 1994	Not feedback
Freeborn 1997	Not RCT
Fretheim 2006	Not core
Furniss 2000	Not feedback
Ganz 2005	Not core

Study	Reason for exclusion
Garrouste-Orgeas 2010	Not core
Gask 1991	Outcome was teaching interviewing skills to medical students; feedback did not include audit
Gerbert 1988	Not RCT
Goderis 2010	Not core
Goldberg 1980	Not audit and feedback
Goldberg 1998	Not core
Grimshaw 1998	Insufficient data on results
Gunn 2003	Not RCT
Hall 2001	Not audit and feedback
Hampshire 1999	Insufficient data on results
Hanlon 1996	Not audit and feedback
Harbarth 2002	Not core
Harewood 2008	Feedback focusing on skills
Hargraves 1996	Not audit and feedback
Harris 2005	Not feedback
Hartlaub 1993	Not RCT
Henderson 1979	Cost only
Hetlevik 1998	Not feedback
Hinchey 2010	Not RCT
Hirsch 2002	Not RCT
Hogg 2008	Not core
Holleman 1996	Not RCT
Horbar 2004	Not core
Horowitz 1996	Not core feedback
Howe 1996	Not core
Hulscher 1997	Not RCT
Ilag 2003	Not core

Study	Reason for exclusion
Jaen 2010	Not feedback
Jans 2001	Not RCT
Johansen 1997	Not audit and feedback
Johnson 1976	Not audit or summery of performence
Jones 1996	Procedural skill
Kafuko 1999	Not clearly randomised trial
Katz 2004	Not core
Kerse 1999	Not core
Kinney 2003	No results
Kirwin 2010	Not feedback
Kroenke 1990	Not RCT
Kuilboer 2006	Not feedback
Labarere 2007	not core feedback
Lafata 2007	Not core
Lassen 1992	Not RCT
Lemelin 2001	Not core
Lenderink 2010	Not feedback
Leviton 1999	Not core
Linn 1980	Not audit and feedback
Luders 2010	Not feedback
Lundborg 1999	Not core
MacCosbe 1985	Not audit and feedback
MacGowan 1996	Not RCT
Madridejos-Mora 2004	Not RCT
Mandel 1985	Missing results
Manfredi 1998	Not core
Manheim 1990	Not core
Manning 1986	Not RCT

Study	Reason for exclusion
Martin 2007	No results reported
Mayefsky 1993	Not randomised
Mazzuca 1988	Not audit and feedback
McDermott 2003	Insufficient data on result
McDonel 1997	Not feedback
McPhee 1989	Insufficient data on result
Meehan 2001	Not RCT
Mertz 2010	Not feedback
Metlay 2007	Not core
Meyer 1991	Not summary of performance
Moongtui 2000	Not RCT
Mourad 2010	No results
Munroe 1997	Not RCT
Myers 2004	Not core
Nattinger 1989	Non-equivalent group design with pre-post measures
Nicolas 1996	Not RCT
North of England1992	Missing results
Nyman 1995	Not feedback
O'Connor 1996	Not RCT
O'Connor 2005	Not core
Ogwal-Okeng 2001	Insufficient data on results
Ornstein 2010	Not core
Ottolini 1998	Not audit and feedback
Overbeek 2010	Not core
Papa 1999	Not feedback
Patel 2010	Not feedback
Payne 1978	Not RCT
Pearson 2001	Not RCT, not feedback

Study	Reason for exclusion
Performance 2006	No results
Peters-Klimm 2008	Not core
Peters-Klimm 2009	Not core
Pfirter 2010	Not feedback
Pit 2007	Not core
Pugh 1989	Not RCT
Putnam 1985	Insufficient data on results
Quilitch 1975	Not RCT
Raisch 1999	Not RCT
Rascati 1996	Not feedback
Reid 1977	Cost only
Restuccia 1982	Intervention did not include audit
Reuther 2010	Not core
Rhew 1999	Not RCT
Rollman 2002	Not audit and feedback
Roski 1998	Not core
Rubenstein 1989	Not feedback on performence
Rubenstein 1999	Not feedback
Sanazaro 1978	Not RCT
Seers 2004	Not core
Shaughnessy 1991	Skills, not clinical performance
Simon 2000	Not summary of performance
Simunovic 2010	Not core
Smeele 1998	Not RCT
Smith 1995	Skills
Spector 1989	Intervention was a federal survey process
Steele 1989	Cost only
Stewart 2005	Not feedback

Study	Reason for exclusion
Strasser 2008	Not feedback (no summary of clinical performance)
Strikwerda 1994	Not feedback
Sundaram 2009	Not RCT
Szczepura 1994	Missing results
Taylor 1997	Not RCT
The SUPPORT 1995	No feedback on performence
Thompson 2000	Not core
Van Bruggen 2008	Not core
Van der Sanden 2005	Not feedback
Van der Weijden 1998	Not core
Velikova 2004	Not audit and feedback
Verstappen 2004 b	No results, cost only
Vinicor 1987	Not core
Walsh 2007	Not core
Watkins 1981	Not RCT
Weingarten 2000	Facilitated relay of clinical information, not feedback on clinical performance
Wells 2000	Not core
Welschen 2004	Not core
White 1995	Not feedback on performence
Wing 1987	Not audit and feedback
Wing 1987 (II)	Not audit and feedback
Winickoff 1985	Not RCT
Winkens 1992	Not RCT
Winkens 1997	Insufficient data on results
Yano 2008	Not core
Young 2002	Not core
Zermansky 2002	Not feedback

Study	Reason for exclusion
Zoutman 2010	inadequate information - abstract only

RCT: randomised controlled trial

Characteristics of studies awaiting assessment *[ordered by study ID]*

Bond 2011

Methods	Design: cluster RCT
Participants	Country: USA Setting: outpatient Specialty: internists N health professionals: (77 dialysis centers) N patients: -
Interventions	feedback versus feedback plus educational outreach
Outcomes	Targeted behaviour: vaccinations Baseline performance: moderate
Notes	

Daley 2011

Methods	Design: cluster RCT
Participants	Country: USA Setting: both Specialty: addiction specialists N health professionals: (103 treatment centers) N patients: -
Interventions	feedback versus feedback plus educational outreach
Outcomes	Targeted behaviour: mix Baseline performance: unclear
Notes	

Guldberg 2011

Methods	Design: cluster RCT

Guldberg 2011 *(Continued)*

Participants	Country: Denmark Setting: outpatinet Specialty: family medicine N health professionals: 158 (86 practices) N patients: 2458
Interventions	feedback versus usual care
Outcomes	Targeted behaviour: mix Baseline performance: unclear
Notes	

Ivers 2010

Methods	Design: cluster RCT
Participants	Country: Canada Setting: outpatient Specialty: family medicine N health professionals: 54 N patients: 5000
Interventions	feedback versus feedback with action plan worksheet
Outcomes	
Notes	Protocol

LaPointe 2006

Methods	Design: cluster RCT
Participants	Country: USA Setting: outpatient Specialty: family medicine/internal medicine N health professionals: 66 (45 practices) N patients: 2717
Interventions	feedback plus provider education plus patient mediated versus patient-specific feedback plus intensive provider education plus extended patient mediated interventions
Outcomes	Targeted behaviour: increase prescribing of beta blockers

LaPointe 2006 *(Continued)*

	Baseline performance: moderate
Notes	patient specific versus aggregate feedback compared, but difficult to disentagle from other aspects

Lopez-Picazo 2011

Methods	Design: Cluster-RCT
Participants	Country: Spain Setting: outpatient Specialty: family medicine N health professionals: 265 N patients: 81,805
Interventions	usual care versus written feedback versus written feedback plus group education versus written feedback plus educational outreach
Outcomes	Targeted behaviour: decrease prescribing of drug-interactions Baseline performance: high
Notes	

Mourad 2011

Methods	Design: cluster-RCT
Participants	Country: Netherlands Setting: outpatient Specialty: obgyn N health professionals: (16 clinics) N patients: 1396
Interventions	feedback versus feedback plus educational outreach plus patient mediated tools
Outcomes	Targeted behaviour: mix of guideline indicators for managment of infertility Baseline performance: mix
Notes	

Palmer 1996

Methods	Design: cluster-RCT
Participants	Country: USA

Palmer 1996 *(Continued)*

	Setting: inpatient
	Specialty: unclear
	N health professionals: (16 clinics)
	N patients: -
Interventions	education versus education versus feedback
Outcomes	Targeted behaviour: mix
	Baseline performance: -
Notes	Same practices as Palmer 1995 but different trial?

Sequist 2010

Methods	Design: cluster-RCT
Participants	Country: USA
	Setting: outpatient
	Specialty: primary care
	N health professionals: 124 (8 clinics)
	N patients: -
Interventions	usual care versus race-stratified feedback
Outcomes	Targeted behaviour: mix
	Baseline performance: -
Notes	primary outcomes at patient-level

Smeets 2010

Methods	Design: cluster-RCT
Participants	Country: Netherlands
	Setting: outpatient
	Specialty: primary care
	N health professionals: 993 (112 groups)
	N patients: 23,433
Interventions	usual care versus patient-specific feedback plus education plus financial incentives
Outcomes	Targeted behaviour: decrease prescriptions of acid-suppresants
	Baseline performance: high

Smeets 2010 *(Continued)*

Notes	may not be best classified as feedback

Williams 2011

Methods	Design: cluster-RCT
Participants	Country: USA Setting: inpatient Specialty: cardiac surgeons N health professionals: (458 hospitals) N patients: 361,328
Interventions	usual care versus feedback plus education, plus standardized orders and patient education
Outcomes	Targeted behavious: increase prescriptions of cardiac medications Baseline performance: moderate to high
Notes	

ADDITIONAL TABLES

Table 1. Description of Included Trials (N = 140)

Study Characteristic	Number	Percent	Intervention Characteristic	Number	Percent
Publication Year			Audit and Feedback alone	49	35.0
2006-2010	32	22.9	Multifaceted intervention with AF as core feature	91	65.0
1996-2005	76	54.3	with Case management or team change	3	2.1
1986-1995	20	14.3	with Clinician education (not outreach)	48	34.3
before 1986	12	8.6	with Educational outreach	28	20.0
Country			with Clinician reminders, including decision support	17	12.1
USA	69	49.3	with Patient intervention (eg. self mgmt/reminders)	8	5.7
UK or Ireland	21	15.0	with Continuous quality improvement	9	6.4

Table 1. Description of Included Trials (N = 140) *(Continued)*

Canada	11	7.9	with Financial incentives	5	3.6
Australia or New Zealand	10	7.1	**Format**		
Other	29	20.7	Verbal	13	9.3
Unit of Allocation			Written	84	60.0
Provider	51	36.4	Both	32	22.9
Many Providers/ Groups	88	62.9	Unclear	11	7.9
Unclear	1	0.7	**Source**		
Unit of Analysis			Supervisor/colleague	13	9.3
Patient	81	57.9	Employer	15	10.7
Provider	29	20.7	Investigators/unclear	112	80.0
Many Providers/ Groups	29	20.7	**Frequency**		
Unclear	1	0.7	Weekly	11	7.9
Risk of Bias			Monthly	19	13.6
Low	45	32.1	Repeated less than monthly	36	25.7
Unclear	70	50.0	Once only	68	48.6
High	25	17.9	**Instructions for Improvement**		
Number of Arms in Trial			Goal-setting	11	7.9
Two	98	70.0	Action planning	41	29.3
Three	22	15.7	Both	4	2.9
Four	20	14.3	Neither	84	60.0
Clinical Setting			**Direction of Change Required**		
Outpatient	94	67.1	Increase current behaviour	57	40.7
Inpatient	36	25.7	Decrease current behaviour	29	20.7
Other/unclear	10	7.1	Mix or unclear	55	39.3
Medical Specialty (could include more than one)			**Targeted Health Professional (could include more than one)**		
GP/Family physician	84	60.0	Physician	121	86.4

Table 1. Description of Included Trials (N = 140) *(Continued)*

Internists	60	42.9	Nurses	16	11.4
Other	40	28.6	Pharmacists	5	3.6
			Other	3	2.1
			Clinical Topic / Targeted Behaviour (could be more than one)		
			Diabetes/Cardiovascular disease management	30	21.4
Size of trial	***Median***	***IQR***	Laboratory testing/radiology	21	15.0
Providers (when providers allocated)	56	28-139	Prescribing	31	22.1
Groups (when many providers allocated)	32	19-69	Other	50	41.4

Table 2. Assessment of Heterogeneity: results from meta regression

Characteristic of the Feedback or Recipient or Trial		**Effect**
Format of feedback		**P = 0.020**
	Verbal	3.38
	Written	9.50
	Both verbal and written	11.23
	Not clear	5.27
Source of feedback		**P < 0.001**
	A supervisor or colleague	16.50
	A 'professionals standards review organization' or employer	2.37
	The investigators	5.04
	Not clear	5.48
Frequency of feedback		**P < 0.001**
	Frequent (up to weekly)	1.44
	Moderate (up to monthly)	9.83
	Infrequent (less than monthly)	4.78
	Once only	2.56

Table 2. Assessment of Heterogeneity: results from meta regression *(Continued)*

	Unclear	18.12
Instructions for improvement		**P < 0.001**
	Explicit, measurable target/goal, but no action plan	2.52
	Action plan	9.57
	Both	11.09
	Neither	6.20
Direction of change required		**P < 0.001**
	Increase current behaviour	4.34
	Decrease current behaviour	10.54
	Change behaviour or mix or unclear	7.16
Baseline performance		**P = 0.007**
	at 25%	9.11
	at 50%	7.07
	at 75%	5.03
Profession of recipient		**P = 0.561**
	Physician	7.90
	Non-physician	6.80
Risk of bias		**P = 0.679**
	Low risk of bias	7.68
	Unclear	7.02
	High risk of bias (not included in primary analysis)	n/a

Table 3. Exploratory analysis: meta regression with targeted behaviour

Characteristic of the Feedback or Recipient or Trial	**Effect**
Type of professional practice	**P < 0.001**
Diabetes/CVD	5.91
Laboratory testing/radiology referrals	4.21
Prescribing	11.11

Table 3. Exploratory analysis: meta regression with targeted behaviour *(Continued)*

	Other	4.71
Format of feedback		**P < 0.001**
	Verbal	2.42
	Written	5.86
	Both verbal and written	10.07
	Not clear	7.60
Source of feedback		**P < 0.001**
	A supervisor or colleague	13.71
	A 'professionals standards review organization' or employer	2.44
	The investigators	4.95
	Not clear	4.85
Frequency of feedback		**P = 0.002**
	Frequent (up to weekly)	3.09
	Moderate (up to monthly)	9.58
	Infrequent (less than monthly)	6.28
	Once only	3.59
	Unclear	9.89
Instructions for improvement		**P < 0.001**
	Explicit, measurable target/goal, but no action plan	2.84
	Action plan	9.30
	Both	7.18
	Neither	6.63
Direction of change required		**P = 0.525**
	Increase current behaviour	6.64
	Decrease current behaviour	7.13
	Change behaviour or mix or unclear	5.70
Baseline performance		**P = 0.002**
	at 25%	8.72

Table 3. Exploratory analysis: meta regression with targeted behaviour *(Continued)*

	at 50%	6.75
	at 75%	4.77
Profession of recipient		**P = 0.059**
	Physician	5.04
	Non-physician	7.94
Risk of bias		**P = 0.454**
	Low risk of bias	5.88
	Unclear	7.09
	High risk of bias (not included in primary analysis)	n/a

APPENDICES

Appendix 1. Selected variables considered for inclusion in meta-regression analysis

Variable	Previous version	Comments	Decision for new version
Intensity of AF	In analysis	Previous approach unhelpful	Remove
Complexity of behavior	In analysis	Not predictive	Remove
Seriousness of outcome	In analysis	Not predictive	Remove
Baseline compliance	In analysis	Predictive	Keep
Risk of bias	In analysis	Update based on revised Cochrane Handbook requirements	Keep
Peer comparison	In analysis	Not predictive	Remove
Close to time of decision making	New	Mugford 1991 Judged to be difficult to abstract	Not added
Quality of data, Motivation of recipients	New	Van der Veer 2010 Based on perception of recipients, thus difficult to abstract	Not added
Organizational support/culture	New	Van der Veer 2010, Hysong 2006 Judged to be difficult to abstract	Not added
Participative intervention	New	Van der Veer 2010, Locke 2002	Not added

(Continued)

		Judged to be difficult to abstract	
Profession of recipient	New	Physicians behaviour is likely harder to change	Add
Direction of change	New	Carlsen 2007. Qualitative evidence that decreasing is harder.	Add
Correct solution information, goal-setting and action- plans	New	Locke 2002, Hysong 2009, Sniehotta 2009, Gardner 2010 Theory suggests these should help	Add
Tailoring of intervention after assessment of barriers	Descriptive	Grimshaw 2004. Not feedback-specific.	Not added
Clinical topic	Descriptive	No clear hypothesis to test	Not added
Setting	Descriptive	Axt-Adam 1993. Likely important, but no clear hypothesis to test	Not added
Frequency	Part of intensity	Hysong 2006 found this to be associated with high performing groups	Keep
Format (written or verbal)	Part of intensity	Hysong 2009 Very important in recent reanalysis	Keep
Source	Part of intensity	Hysong 2006 and other qualitative work suggest that trust matters	Keep
Recipient	Part of intensity	Judged to be less important than other aspects related to intensity	Remove
Setting (inpatient versus outpatient)	New	Inpatient feedback may be more effective given more resources and often higher acuity of target/problem	Add

Appendix 2. Electronic Search Strategies

CENTRAL

#1	MeSH descriptor Clinical Audit, this term only	5
#2	MeSH descriptor Medical Audit, this term only	316
#3	MeSH descriptor Nursing Audit, this term only	58
#4	MeSH descriptor Dental Audit, this term only	4
#5	MeSH descriptor Management Audit, this term only	8
#6	MeSH descriptor Benchmarking, this term only	120
#7	MeSH descriptor Commission on Professional and Hospital Activities, this term only	4

(Continued)

#8	MeSH descriptor Feedback, this term only	799
#9	MeSH descriptor Feedback, Psychological, this term only	179
#10	MeSH descriptor Utilization Review, this term only	262
#11	MeSH descriptor Drug Utilization Review, this term only	218
#12	MeSH descriptor Concurrent Review, this term only	5
#13	MeSH descriptor Peer Review, Health Care, this term only	29
#14	(audit or audits or auditing or feedback or benchmark*):ti,ab	4215
#15	(review NEAR/3 record* or chart NEXT review or practice NEXT data or hospital* NEXT data):ti,ab	1692
#16	(#1 OR #2 OR #3 OR #4 OR #5 OR #6 OR #7 OR #8 OR #9 OR #10 OR #11 OR #12 OR #13 OR #14 OR #15)	6861
#17	MeSH descriptor Health Personnel explode all trees	4673
#18	MeSH descriptor Hospitals explode all trees	3187
#19	MeSH descriptor Professional Practice explode all trees	3354
#20	MeSH descriptor Family Practice, this term only	2201
#21	MeSH descriptor Professional Competence, this term only	139
#22	MeSH descriptor Clinical Competence, this term only	1312
#23	MeSH descriptor Physician's Practice Patterns, this term only	1180
#24	MeSH descriptor Nurse's Practice Patterns, this term only	7
#25	MeSH descriptor Dentist's Practice Patterns, this term only	21
#26	MeSH descriptor Quality Assurance, Health Care, this term only	735
#27	MeSH descriptor Quality of Health Care, this term only	844
#28	(health* NEXT personnel or "health care personnel" or physician* or doctor* or clinician* or nurse* or provider* or practitioner* or resident* or professional* or nursing or clinical) NEAR/3 (skill or skills or behaviour or behavior or competence):ti,ab	1788
#29	(clinical or medical or dental or private or general or family or professional or hospital*) NEXT practice*:ti,ab	8108
#30	(practice NEAR/2 pattern*):ti,ab	186
#31	quality NEXT (assurance or improvement or control):ti,ab	1106
#32	(health* or care) NEAR/2 quality:ti,ab	3771

(Continued)

#33	performance:ti,ab	25218
#34	(influenc* NEAR/3 behaviour* or influenc* NEAR/3 behavior* or chang* NEAR/3 behaviour* or chang* NEAR/3 behavior*):ti,ab	2560
#35	(#17 OR #18 OR #19 OR #20 OR #21 OR #22 OR #23 OR #24 OR #25 OR #26 OR #27 OR #28 OR #29 OR #30 OR #31 OR #32 OR #33 OR #34)	51494
#36	(#16 AND #35)	2344
#37	audit* NEAR/3 feedback:ti,ab	190
#38	(#36 OR #37)	2431

MEDLINE

1.	(audit* adj3 feedback).tw.	1268
2.	Clinical Audit/	327
3.	Medical Audit/	13185
4.	Nursing Audit/	2838
5.	Dental Audit/	278
6.	Management Audit/	2272
7.	Benchmarking/	8052
8.	"Commission on Professional and Hospital Activities"/	189
9.	Feedback/	24251
10.	Feedback, Psychological/	1457
11.	Utilization Review/	6467
12.	Drug Utilization Review/	2570
13.	Concurrent Review/	372
14.	Peer Review, Health Care/	1167
15.	(audit or audits or auditing).tw.	19672
16.	feedback.tw.	58889
17.	(review adj3 record?).tw.	7594
18.	chart review.tw.	14203

(Continued)

19.	(practice data or hospital* data).tw.	2593
20.	benchmark*.tw.	9736
21.	or/2-20	145800
22.	exp Health Personnel/	318705
23.	exp Hospitals/	174434
24.	exp Professional Practice/	197907
25.	Family Practice/	59744
26.	Professional Competence/	16987
27.	Clinical Competence/	51638
28.	Physician's Practice Patterns/	32063
29.	Nurse's Practice Patterns/	130
30.	Dentist's Practice Patterns/	1306
31.	Quality Assurance, Health Care/	42446
32.	Quality of Health Care/	47308
33.	((health* personnel or health care personnel or physician? or doctor? or clinician? or nurse? or provider? or practitioner? or resident? or professional? or nursing or clinical) adj3 (skill or skills or behaviour or behavior or competence)).tw.	21876
34.	((clinical or medical or dental or private or general or family or professional or hospital?) adj practice?).tw.	128340
35.	(practice pattern? or pattern of practice).tw.	3480
36.	(quality adj (assurance or improvement or control)).tw.	43055
37.	(health care quality or healthcare quality or quality of healthcare or quality of health care or quality of care).tw.	25773
38.	performance.tw.	367757
39.	((influenc* or chang*) adj3 (behaviour* or behavior*)).tw.	36099
40.	or/22-39	1282788
41.	randomized controlled trial.pt.	307057
42.	controlled clinical trial.pt.	83492
43.	(randomi* or randomly).tw.	402940
44.	or/41-43	572824

(Continued)

45.	Animals/	4756026
46.	Humans/	11642321
47.	45 not (45 and 46)	3521849
48.	44 not 47	525534
49.	21 and 40 and 48	2920
50.	1 and 48	166
51.	49 or 50	2975
52.	(2005* or 2006* or 2007* or 2008* or 2009* or 2010*).ed,ep,yr.	4235977
53.	51 and 52	1380

EMBASE

1.	(audit* adj3 feedback).tw.	1378
2.	Medical Audit/	21134
3.	Feedback System/	37936
4.	Negative Feedback/	6456
5.	Positive Feedback/	2913
6.	"Utilization Review"/	56364
7.	"Medical Record Review"/	18847
8.	(audit or audits or auditing).tw.	25835
9.	feedback.tw.	65004
10.	(review adj3 record?).tw.	8496
11.	chart review.tw.	17507
12.	(practice data or hospital* data).tw.	3017
13.	benchmark*.tw.	12228
14.	or/2-13	223611
15.	exp Health Care Personnel/	608667
16.	exp Hospital/	413165

(Continued)

17.	exp Professional Practice/	207429
18.	Professional Competence/	16932
19.	Nursing Competence/	291
20.	Clinical Competence/	32803
21.	Health Care Quality/	142112
22.	Quality Control/	87850
23.	((health* personnel or health care personnel or physician? or doctor? or clinician? or nurse? or provider? or practitioner? or resident? or professional? or nursing or clinical) adj3 (skill or skills or behaviour or behavior or competence)).tw.	24921
24.	((clinical or medical or dental or private or general or family or professional or hospital?) adj practice?).tw.	154741
25.	(practice pattern? or pattern of practice).tw.	4014
26.	(quality adj (assurance or improvement or control)).tw.	54854
27.	(health care quality or healthcare quality or quality of healthcare or quality of health care or quality of care).tw.	30068
28.	performance.tw.	439514
29.	((influenc* or chang*) adj3 (behaviour* or behavior*)).tw.	41092
30.	or/15-29	1806950
31.	Randomized Controlled Trial/	285934
32.	(randomi* or randomly).tw.	491780
33.	or/31-32	570426
34.	Nonhuman/	3542502
35.	33 not 34	517018
36.	14 and 30 and 35	3872
37.	1 and 35	163
38.	36 or 37	3918
39.	38 not medlinex00ae.cr.	2612
40.	2010*.em.	1107279
41.	39 and 40	452

CINAHL

S47	S46 - Limiters - Exclude MEDLINE records	268
S46	S44 or S45	1104
S45	S42 and S43	83
S44	S13 and S36 and S42	1079
S43	TI (audit* and feedback) or AB (audit* and feedback)	482
S42	S37 or S38 or S39 or S40 or S41	130326
S41	TI ((randomi* or randomly)) or AB ((randomi* or randomly))	69304
S40	(MH "Simple Random Sample")	272
S39	(MH "Random Sample")	16480
S38	(MH "Random Assignment")	24528
S37	(MH "Clinical Trials")	69429
S36	S14 or S15 or S16 or S17 or S18 or S19 or S20 or S21 or S22 or S23 or S24 or S25 or S26 or S27 or S28 or S29 or S30 or S31 or S32 or S33 or S34 or S35	464891
S35	TI (influenc* N3 behaviour* or influenc* N3 behavior* or chang* N3 behaviour* or chang* N3 behavior*) or AB (influenc* N3 behaviour* or influenc* N3 behavior* or chang* N3 behaviour* or chang* N3 behavior*)	8481
S34	TI performance or AB performance	45340
S33	TI ("health care quality" or "healthcare quality" or quality W1 healthcare or quality W2 care) or AB ("health care quality" or "healthcare quality" or quality W1 healthcare or quality W2 care)	15231
S32	TI (quality W0 assurance or quality W0 improvement or quality W0 control) or AB (quality W0 assurance or quality W0 improvement or quality W0 control)	8532
S31	TI practice N1 pattern* or AB practice N1 pattern*	970
S30	TI (clinical W0 practice* or medical W0 practice* or dental W0 practice* or private W0 practice* or general W0 practice* or family W0 practice* or professional W0 practice* or hospital* W0 practice*) or AB (clinical W0 practice* or medical W0 practice* or dental W0 practice* or private W0 practice* or general W0 practice* or family W0 practice* or professional W0 practice* or hospital* W0 practice*)	30978
S29	TI ("health personnel" N3 competence or "healthcare personnel" N3 competence or "health care personnel" N3 competence or physician N3 competence or physicians N3 competence or doctor N3 competence or doctors N3 competence or clinician N3 competence or clinicians N3 competence or nurse N3 competence or nurses N3 competence or provider N3 competence or providers N3 competence or practitioner N3 competence or practitioners N3	2237

(Continued)

	competence or resident N3 competence or residents N3 competence or professional N3 competence or professionals N3 competence or nursing N3 competence or clinical N3 competence) or AB ("health personnel" N3 competence or "healthcare personnel" N3 competence or "health care personnel" N3 competence or physician N3 competence or physicians N3 competence or doctor N3 competence or doctors N3 competence or clinician N3 competence or clinicians N3 competence or nurse N3 competence or nurses N3 competence or provider N3 competence or providers N3 competence or practitioner N3 competence or practitioners N3 competence or resident N3 competence or residents N3 competence or professional N3 competence or professionals N3 competence or nursing N3 competence or clinical N3 competence)	
S28	TI ("health personnel" N3 behavior or "healthcare personnel" N3 behavior or "health care personnel" N3 behavior or physician N3 behavior or physicians N3 behavior or doctor N3 behavior or doctors N3 behavior or clinician N3 behavior or clinicians N3 behavior or nurse N3 behavior or nurses N3 behavior or provider N3 behavior or providers N3 behavior or practitioner N3 behavior or practitioners N3 behavior or resident N3 behavior or residents N3 behavior or professional N3 behavior or professionals N3 behavior or nursing N3 behavior or clinical N3 behavior) or AB ("health personnel" N3 behavior or "healthcare personnel" N3 behavior or "health care personnel" N3 behavior or physician N3 behavior or physicians N3 behavior or doctor N3 behavior or doctors N3 behavior or clinician N3 behavior or clinicians N3 behavior or nurse N3 behavior or nurses N3 behavior or provider N3 behavior or providers N3 behavior or practitioner N3 behavior or practitioners N3 behavior or resident N3 behavior or residents N3 behavior or professional N3 behavior or professionals N3 behavior or nursing N3 behavior or clinical N3 behavior)	1840
S27	TI ("health personnel" N3 behaviour or "healthcare personnel" N3 behaviour or "health care personnel" N3 behaviour or physician N3 behaviour or physicians N3 behaviour or doctor N3 behaviour or doctors N3 behaviour or clinician N3 behaviour or clinicians N3 behaviour or nurse N3 behaviour or nurses N3 behaviour or provider N3 behaviour or providers N3 behaviour or practitioner N3 behaviour or practitioners N3 behaviour or resident N3 behaviour or residents N3 behaviour or professional N3 behaviour or professionals N3 behaviour or nursing N3 behaviour or clinical N3 behaviour) or AB ("health personnel" N3 behaviour or "healthcare personnel" N3 behaviour or "health care personnel" N3 behaviour or physician N3 behaviour or physicians N3 behaviour or doctor N3 behaviour or doctors N3 behaviour or clinician N3 behaviour or clinicians N3 behaviour or nurse N3 behaviour or nurses N3 behaviour or provider N3 behaviour or providers N3 behaviour or practitioner N3 behaviour or practitioners N3 behaviour or resident N3 behaviour or residents N3 behaviour or professional N3 behaviour or professionals N3 behaviour or nursing N3 behaviour or clinical N3 behaviour)	904
S26	TI ("health personnel" N3 skills or "healthcare personnel" N3 skills or "health care personnel" N3 skills or physician N3 skills or physicians N3 skills or doctor N3 skills or doctors N3 skills or clinician N3 skills or clinicians N3 skills or nurse N3 skills or nurses N3 skills or provider N3 skills or providers N3 skills or practitioner N3 skills or practitioners N3 skills or resident N3 skills or residents N3 skills or professional N3 skills or professionals N3 skills or nursing N3 skills or clinical N3 skills) or AB ("health personnel" N3 skills or "healthcare personnel" N3 skills or "health care personnel" N3 skills or physician N3 skills or physicians N3 skills or doctor N3 skills or doctors N3 skills or clinician N3 skills or clinicians N3 skills or nurse N3 skills or nurses N3 skills or provider N3 skills or providers N3 skills or practitioner N3 skills or practitioners N3 skills or resident N3 skills or residents N3 skills or professional N3 skills or professionals N3 skills or nursing N3 skills or clinical N3 skills)	6585

(Continued)

S25	TI ("health personnel" N3 skill or "healthcare personnel" N3 skill or "health care personnel" N3 skill or physician N3 skill or physicians N3 skill or doctor N3 skill or doctors N3 skill or clinician N3 skill or clinicians N3 skill or nurse N3 skill or nurses N3 skill or provider N3 skill or providers N3 skill or practitioner N3 skill or practitioners N3 skill or resident N3 skill or residents N3 skill or professional N3 skill or professionals N3 skill or nursing N3 skill or clinical N3 skill) or AB ("health personnel" N3 skill or "healthcare personnel" N3 skill or "health care personnel" N3 skill or physician N3 skill or physicians N3 skill or doctor N3 skill or doctors N3 skill or clinician N3 skill or clinicians N3 skill or nurse N3 skill or nurses N3 skill or provider N3 skill or providers N3 skill or practitioner N3 skill or practitioners N3 skill or resident N3 skill or residents N3 skill or professional N3 skill or professionals N3 skill or nursing N3 skill or clinical N3 skill)	1090
S24	(MH "Quality of Nursing Care")	5823
S23	(MH "Quality of Health Care")	25796
S22	(MH "Quality Assurance")	9381
S21	(MH "Prescribing Patterns")	896
S20	(MH "Practice Patterns")	2424
S19	(MH "Nursing Skills")	2010
S18	(MH "Clinical Competence")	13517
S17	(MH "Professional Competence")	6233
S16	(MH "Professional Practice+")	105318
S15	(MH "Hospitals+")	46203
S14	(MH "Health Personnel+")	239995
S13	S1 or S2 or S3 or S4 or S5 or S6 or S7 or S8 or S9 or S10 or S11 or S12	29528
S12	TI benchmark* or AB benchmark*	2785
S11	TI hospital* W0 data or AB hospital* W0 data	525
S10	TI "practice data" or AB "practice data"	251
S9	TI "chart review" or AB "chart review"	3089
S8	TI review N3 record* or AB review N3 record*	1984
S7	TI feedback or AB feedback	7593
S6	TI (audit or audits or auditing or feedback) or AB (audit or audits or auditing or feedback)	14118
S5	(MH "Utilization Review")	962
S4	(MH "Feedback")	2845
S3	(MH "Benchmarking")	3141

(Continued)

S2	(MH "Nursing Audit")	612
S1	(MH "Audit")	6010

Reported search process/search strategies in previous versions of the review

1. Jamtvedt 2003

Jamtvedt G, Young JM, Kristoffersen DT, Thomson O'Brien MA, Oxman AD. Audit and feedback: effects on professional practice and health care outcomes. *The Cochrane Database of Systematic Reviews* 2003, Issue 3. Art. No.: CD000259. DOI: 10.1002/14651858.CD000259.

Search methods for identification of studies

The review has been updated primarily by using the EPOC register and pending file. We identified all articles in the Cochrane Effective Practice and Organisation of Care (EPOC) register in January 2001 that had been coded as an RCT or clinical controlled trial (CCT) and as 'audit and feedback'. The EPOC pending file (studies selected from the EPOC search strategy results and awaiting assessment) was also searched in January 2001 using the terms 'audit' or 'feedback'. In addition the previous MEDLINE strategy was used to search MEDLINE from January 1997 to April 2000 and any articles already identified by the EPOC strategy were excluded. This search did not generate any relevant additional articles and therefore was not repeated. The reference lists of new articles that were obtained were reviewed.

Previous searches built upon earlier reviews (Thomson 1995; Davis 1995; Oxman 1995; Davis 1992). We searched MEDLINE from January 1966 to June 1997 without language restrictions. These search terms were used: explode education, professional (tw), explode quality of health care, chart review: or quality assurance (tw), feedback (sh), audit (tw,sh) combined with these methodolological terms: clinical trial (pt), random allocation (sh), randomised controlled trials (sh), double-blind method (sh), single-blind method (sh), placebos (sh), all random: (tw). The Research and Development Resource Base in Continuing Medical Education (RDRB/CME) (Davis 1991) was also searched. The reference lists of related systematic reviews and all articles obtained were reviewed.

An updated search was done in November 2002. Potentially relevant studies found with the updated search are included under References to studies awaiting assessment.

2. Jamtvedt 2006

Jamtvedt G, Young JM, Kristoffersen DT, O'Brien MA, Oxman AD. Audit and feedback: effects on professional practice and health care outcomes. *Cochrane Database of Systematic Reviews* 2006, Issue 2. Art. No.: CD000259. DOI: 10.1002/14651858.CD000259.pub2.

Search methods for identification of studies

The review has been updated primarily by using the EPOC register and pending file. We identified all articles in the Cochrane Effective Practice and Organisation of Care (EPOC) register in January 2004 that had been coded as an RCT or clinical controlled trial (CCT) and as 'audit and feedback'. The EPOC pending file (studies selected from the EPOC search strategy results and awaiting assessment) was also searched in January 2004 using the terms 'audit' or 'feedback'. In addition the previous MEDLINE strategy was used to search MEDLINE from January 1997 to April 2000 and any articles already identified by the EPOC strategy were excluded. This search did not generate any relevant additional articles and therefore was not repeated. The reference lists of new articles that were obtained were reviewed.

Previous searches built upon earlier reviews (Thomson 1995; Davis 1995; Oxman 1995; Davis 1992). We searched MEDLINE from January 1966 to June 1997 without language restrictions. These search terms were used: explode education, professional (non sh), explode quality of health care, chart review: or quality assurance (tw), feedback (sh), audit (tw,sh) combined with these methodolological terms: clinical trial (pt), random allocation (sh), randomised controlled trials (sh), double-blind method (sh), single-blind method (sh), placebos (sh), all random: (tw). The Research and Development Resource Base in Continuing Medical Education (RDRB/CME) (Davis 1991) was also searched. The reference lists of related systematic reviews and all articles obtained were reviewed.

An updated search was done in February 2006. Potentially relevant studies are included under References to studies awaiting assessment.

WHAT'S NEW

Date	Event	Description
5 June 2012	Amended	Risk of bias tables updated

HISTORY

Protocol first published: Issue 3, 1996
Review first published: Issue 1, 1998

Date	Event	Description
16 May 2012	New search has been performed	New search, 32 additional studies included.
16 May 2012	New citation required and conclusions have changed	32 new studies, new authors on team.
30 September 2011	New search has been performed	Identified studies awaiting assessment
10 December 2010	Amended	Updated search applied for revised protocol
8 November 2010	Amended	further edits to protocol
8 November 2010	Amended	edits to protocol
29 April 2008	Amended	Converted to new review format.

CONTRIBUTIONS OF AUTHORS

NI, GJ, and ADO prepared the protocol. NI, GJ, SFl, SFr, and JY applied the inclusion criteria, assessed the quality and extracted the data for the included studies. JOJ conducted the quantitative analyses. NI, GJ, SFl, and ADO conducted the qualitative analyses. NI drafted the manuscript with input from GJ and ADO. All authors provided comments on the manuscript. MJ conducted searches for the literature.

DECLARATIONS OF INTEREST

JMG and ADO are editors in the Cochrane Effective Practice and Organisation of Care group - the peer review process for this review was undertaken independently by another editor. Authors of included trials were not involved in the assessment or data abstraction from these trials.

SOURCES OF SUPPORT

Internal sources

- Norwegian Knowledge Centre for the Health Services, Norway.
- Surgical Outcomes Research Centre, Central Sydney Area Health Service, Australia.
- Needs Assessment & Health Outcome Unit, Central Sydney Area Health Service, Australia.
- Hamilton Regional Cancer Centre, Canada.
- Ottawa Hospital Research Institute, Canada.
- University of Ottawa, Canada.
- Department of Family Practice, Womens College Hospital, Canada.
- Department of Family and Community Medicine (DFCM), University of Toronto, Canada.

NI is supported by a DFCM Doctoral Fellowship Award

External sources

- Canadian Institutes for Health Research (CIHR), Canada.

NI is supported by a CIHR Fellowship

- National Health and Medical Research Council (NHMRC), Australia.

SF is supported by an NHMRC post-doctoral Fellowship

- Canada Research Foundation, Canada.

  JG holds a Canada Research Chair in Health Knowledge Transfer and Uptake

INDEX TERMS

Medical Subject Headings (MeSH)

*Feedback, Psychological; Education, Medical, Continuing; Health Personnel [standards]; Health Services Research; Medical Audit [*standards]; Outcome Assessment, Health Care; Practice Patterns, Physicians' [*standards]; Professional Practice [*standards]; Randomized Controlled Trials as Topic

MeSH check words

Humans