

Previous versions of the review titled **Audit and feedback: effects on professional practice and health care outcomes** were published in the *Cochrane Database of Systematic Reviews* in the following years:

- 2000:** Thomson O'Brien MA, Oxman AD, Davis DA, Haynes RB, Freemantle N, Harvey EL. PMID: 10796520
2003: Jamtvedt G, Young JM, Kristoffersen DT, Thomson O'Brien MA, Oxman AD. PMID: 12917891
2006: Jamtvedt G, Young JM, Kristoffersen DT, O'Brien MA, Oxman AD. PMID: 16625533
2012: Ivers N, Jamtvedt G, Flottorp S, Young JM, Odgaard-Jensen J, French SD, O'Brien MA, Johansen M, Grimshaw J, Oxman AD. PMID: 22696318

Rationale for the update:

- The latest version of the review¹ has been cited 1920 times and used in 14 guidelines, suggesting audit and feedback (A&F) continues to be of significant interest to healthcare providers and researchers.
- The 2012 publication included data from 140 studies, however over 100 new relevant trials have been published since 2012 which could have a meaningful impact on the review findings.
- Developments in frameworks related to A&F have occurred to understand key factors related to effectiveness.
- This update will expand the scope and methods used in the previous reviews to explore specific factors that explain/optimize the effectiveness of A&F.

Summary of main differences between 2012 review and planned update:

	<i>Previous review(s):</i>	<i>Current update:</i>
Eligible A&F interventions	Only included studies if A&F was the core intervention in at least one arm	Will also include studies if A&F is a component of a multicomponent intervention (more inclusive)
Eligible outcomes	Included studies reporting health provider process outcomes and/or patient outcomes	Will only include studies investigating process outcomes that impact health provider behaviour as this is more directly attributable to the intervention and less heterogenous
Identification of intervention components	Not done in previous reviews	Will identify and categorize the components of each intervention arm using two established coding schemes and will extract details about the nature of the A&F intervention using a theory-informed set of variables
Analysis comparisons	Compared guidance-concordant behaviour for arms with A&F vs arms without A&F	Will use arm-level analysis to include all data from studies featuring A&F in multiple arms and to examine the effects of various factors on effectiveness of A&F for guideline-concordant behaviours of health professionals (as feasible)

Title: Audit and feedback: effects on professional practice

Primary investigator: Dr. Noah Ivers

Current status of review: Data extraction ongoing

Projected date of completion: July 2022

Audit and feedback (A&F): Defined as a ‘summary of clinical performance over a specified period of time’. The audit component involves measuring quality of care against standards of practice (i.e., guideline recommendations) using chart audits, electronic health records, patient surveys and/or routine administrative data. The feedback component involves providing a report on performance to relevant healthcare professionals, administrators or organizations with the purpose of prompting change and may be delivered in a written, electronic or verbal format. A&F may also be described as clinical performance feedback, benchmarking, report cards, dashboards, practice profiles, etc.

Review question(s): What is the effectiveness of A&F for health professional practices? Which factors (targeted professionals, targeted clinical behaviors, contexts, characteristics of the A&F itself, co-interventions) are associated with greater effects across A&F trials?

Literature search: Literature searches conducted in MEDLINE (Ovid), EMBASE (Ovid), CINAHL (Ebsco), the Cochrane Library, and clinicaltrials.gov from June 2010 to June 2020. WHO ICTRP was only searched until February 2019; no information was available beyond that due to the COVID-19 pandemic.

Study selection: For studies newly identified by the literature searches, two reviewers independently screened titles and abstracts, followed by full text articles. All discrepancies were resolved through discussion. For training, all reviewers independently screened 50 randomly selected citations and discussed disagreements until 95% agreement was reached. Studies included in the previous review were reassessed to account for changes in the eligibility criteria and methods for this updated review.

Studies were screened for inclusion using the following eligibility criteria:

Participants: Healthcare professionals directly involved in patient care. Healthcare professionals in postgraduate training were included, but studies involving only undergraduate students were not.

Interventions: A&F alone or as part of a multicomponent intervention in at least one arm.

Comparator: Other quality improvement (QI) intervention, no intervention or usual care.

Outcomes: Objectively measured health professional practice outcomes (e.g., prescribing, test-ordering, counselling, referrals, etc.). Excluded studies that only measure patient health outcomes, knowledge/attitudes, or performance in a test situation.

Study types: Randomized trials including cluster and step-wedge trials. To qualify as a valid cluster-trial, there must be at least 4 clusters per experimental treatment.

Data Extraction: A pre-determined data collection form will be used to collect information on study design, participants, setting, methods, type of interventions, type of targeted clinical behavior, process outcomes, and results. Informed by recent advances in theory regarding best practices for implementing A&F, significant changes were made to the variables to be extracted describing the intervention arms, such that all studies will be re-extracted. Data will be extracted independently by two review authors. Discrepancies between authors will be resolved through discussion.

We will extract all outcomes assessing changes in desired professional behavior (which we assume to capture a latent concept of increasing guideline-concordant clinical practice). We expect several such outcomes to be reported in each study with no straightforward way to prioritize them. To capture the uncertainty induced by the choice of outcome in the analysis, we will use a bootstrap-based approach (see analysis section below).

Where the outcome is measured at multiple points in time, we will take the timepoint most proximal to the end of the intervention. Data for cross-over trials will be taken from the final time point before cross-over and data from step-wedge trials will not be included in the meta-regression. For multi-arm trials, we will identify the least-intensive-intervention/no intervention arm and use this arm as the control for other comparisons (dividing the sample equally for the number of pairwise comparisons).

Components of each A&F intervention that align with our theoretical understanding of intervention mechanisms and offer face validity in terms of the steps that intervention developers must consider for implementation: 1) *audit* (number of indicators², aggregate vs patient-level data³); 2) *feedback* (format⁴, source⁵, type of comparator⁶, frequency⁷); and 3) *facilitation* (action planning⁸, participant involvement) will be extracted.

In addition, the team will use the following two methods to extract the components of each of the study interventions in each study arm (in duplicate), which was not done in previous versions of the review:

Behaviour Change Techniques (BCTs)⁹: Trained extractors developed a codebook and will apply the established BCT taxonomy (version 1) describing 93 techniques to change behavior.

Effective Practice and Organisation of Care (EPOC) framework¹⁰: Trained extractors will code based on the established EPOC classification scheme describing 16 broad categories of health systems interventions.

Risk of bias assessment: Two reviewers will independently assess the risk of bias for each newly included study using the criteria outlined in the Cochrane Handbook¹¹. Discrepancies will be resolved by discussion or by involving a third reviewer as needed.

Data synthesis and analysis:

We will organize and summarize the data to assess the following comparisons for the main latent outcome of 'guideline concordant care': *i) A&F with or without co-interventions vs usual care; ii) A&F alone vs usual care; iii) A&F plus co-intervention vs usual care; iv) A&F plus co-intervention vs A&F alone; and v) head-to-head trials of different forms of A&F.* We will conduct these analyses for dichotomous outcomes since this is the most common way that outcomes of interest are expressed. We will repeat the meta-analyses of the first comparator group (A&F with or without co-intervention vs usual care comparison) for subjects of outcomes that fall under commonly targeted health professional behaviors of A&F (prescribing and test-ordering).

To select outcome data for meta-analyses from each study, rather than choosing the median as we did in the prior review, we will use a hierarchical bootstrap approach to conduct meta-analyses (sampling studies with replacement, and then, conditional on the sampled studies, sampling outcomes with replacement). We will use standard methods recommended by Cochrane to identify and correct unit of analysis errors and will make the required adjustments before including data from such studies in the meta-analysis (or meta-regression) models. This will enable the inclusion of all outcome data and avoid arbitrary selection of outcomes (e.g., primary outcome).

An equally important objective to updating the overall estimates of effect for A&F is the aim of unpacking the characteristics of interventions associated with greater and lesser effects. We will describe the composition of these interventions and the frequency with which different versions of the interventions have been evaluated in the trials.

Data permitting (and in the presence of a meaningful change observed in the main meta-analyses), we will fit arm-based meta-regression models to explore the association of components of A&F intervention with the overall outcome of guideline concordant care. As with the meta-analyses, we will repeat these meta-regressions for A&F targeting prescribing and test-ordering. Assuming adequate data, the meta-regressions will explore the following A&F characteristics: number of indicators, aggregate vs patient-level data, format, frequency, and source of the feedback and type of comparator used, presence of action plans, and participant involvement in action plans, as well as number of co-interventions (defined using the EPOC taxonomy).

We hypothesize that greater effects of A&F will be observed with:

- Specific features of the audit
 - Only one targeted behaviour (e.g., only prescribing) in the audit because it limits cognitive load and enables focused individual and organizational action.
 - Disaggregated data (recipient's own performance, rather than team) in the audit because it is easier to validate by the recipient and enables the recipient to feel accountable for the results
- Specific features of the feedback
 - Multi-modal feedback format (verbal plus written) because the data may become more memorable this way
 - Delivery of the feedback by a local champion because this may improve credibility
 - Only one (positive or normative) comparator or target rather than multiple to minimize cognitive load and encourage improvement against a goal
 - Personal historical data showing trends in performance because this may improve perceived validity and accountability for improvement

- Repeated delivery of the feedback as this reinforces the feedback loop compared to once-only feedback
- Specific features to enable facilitation
 - Action plans provided so that recipients know specifically what to do to achieve higher scores
 - Co-development of action plans because this may help recipients feel the approach is actionable and social construction of the feedback loop may enable a culture of data-driven quality improvement.

Findings will be summarized in a table along with a rating of the certainty of evidence for the results using the GRADE approach¹².

Dissemination: The full review findings will be published in the Cochrane Database of Systematic Reviews, presented at relevant in-person and online conferences, and a one-page summary of the findings and lay summary will be made available for distribution.

References:

1. Ivers N, Jamtvedt G, Flottorp S, Young J, Odgaard-Jensen J, French S, et al. Audit and feedback: effects on professional practice and healthcare outcomes. *The Cochrane Database Syst Rev.* 2012;6:CD000259. doi: 10.1002/14651858.CD000259.pub3.
2. Hibbard J, Peters E, Slovic P, Finucane M, Tusler M. Making health care quality reports easier to use. *Jt Comm J Qual Improv.* 2001;27(11):591-604. doi: 10.1016/s1070-3241(01)27051-5.
3. Rat C, Pogu C, Le Donne D, Latour C, Bianco G, Nanin F, et al. Effect of physician notification regarding nonadherence to colorectal cancer screening on patient participation in fecal immunochemical test cancer screening: A randomized clinical trial. *JAMA.* 1999;318(9):816-24. doi: 10.1001/jama.2017.11387.
4. Tuti T, Nzinga J, Njoroge M, Brown B, Peek N, English M, et al. A systematic review of electronic audit and feedback: intervention effectiveness and use of behaviour change theory. *Implementation Sci.* 2017;12(1):61. doi.org/10.1186/s13012-017-0590-z.
5. Brown B, Gude WT, Blakeman T, van der Veer SN, Ivers N, Francis JJ, et al. Clinical Performance Feedback Intervention Theory (CP-FIT): a new theory for designing, implementing, and evaluating feedback in health care based on a systematic review and meta-synthesis of qualitative research. *Implementation Sci.* 2019; 14:40. doi.org/10.1186/s13012-019-0883-5.
6. Kiefe C, Allison J, Williams O, Person S, Weaver M, Weissman N. Improving quality improvement using achievable benchmarks for physician feedback: a randomized controlled trial. *JAMA.* 2001;285(22):2871-9. doi: 10.1001/jama.285.22.2871.
7. Lam C, DeRue D, Karam E, Hollenbeck J. The impact of feedback frequency on learning and task performance: Challenging the “more is better” assumption. *Organ Behav Hum Dec.* 2011;116(2):217-28. doi: 10.1016/j.obhdp.2011.05.002
8. Dillman D, Smyth J, Christian L. *Internet, phone, mail, and mixed-mode surveys: The tailored design method.* Hoboken: Wiley; 2014.
9. Michie S, Wood C, Johnston M, Abraham C, Francis J, Hardeman W. Behaviour change techniques: the development and evaluation of a taxonomic method for reporting and describing behaviour change interventions (a suite of five studies involving consensus methods, randomised controlled trials and analysis of qualitative data). *Health Technol Assess.* 2015;19(99):1-188. doi: 10.3310/hta19990.
10. Effective Practice and Organisation of Care (EPOC). EPOC Taxonomy; 2015. epoc.cochrane.org/epoc-taxonomy (accessed 03 Nov 2020).
11. Higgins J P T, Altman D G, Gøtzsche P C, Jüni P, Moher D, Oxman A D et al. The Cochrane Collaboration’s tool for assessing risk of bias in randomised trials. *BMJ.* 2011; 343: d5928. doi:10.1136/bmj.d5928.
12. Schünemann, Holger; Brożek, Jan; Guyatt, Gordon; Oxman, Andrew, editors. *Handbook for grading the quality of evidence and the strength of recommendations using the GRADE approach.* Updated October 2013. Available from <https://gdt.gradeapro.org/app/handbook/handbook.html>.