When is audit and feedback effective in dementia care? A systematic review

Michael J. Sykes*, Jennifer McAnuff, Niina Kolehmainen

Institute for Health and Society, Newcastle University, Richardson Road, Newcastle-Upon-Tyne NE2 4AF, UK

ABSTRACT

Background: Evidence-based care for people with dementia is a priority for patients, carers and clinicians and a policy priority. There is evidence that people with dementia do not always receive such care. Audit and feedback, also known as clinical audit, is an extensively-used intervention to improve care. However, there is uncertainty about the best way to use it.
Objectives: To investigate whether audit and feedback is effective for improving health professionals' care of people with dementia. To investigate whether the content and delivery of audit and feedback affects its effectiveness in the context of health professionals' care for people with dementia.
Design: Systematic review
Data sources: The Cochrane Central Register of Controlled Trials, Prospero, Medline (1946–December week 1 2016), PsycInfo (1967–January 2017), Cinahl (1982–January 2017), HMIC (1979–January 2017), Embase (1974–2017 week 1) databases and the Science Citation Index and Social Science Citation Index were searched combining terms for audit and feedback, health personnel, and dementia.
Review methods: Following screening, the data were extracted using the Template for Intervention Description and Replication (TIDieR), and synthesised graphically using harvest plots and narratively.
Results: Thirteen studies met the inclusion criteria. Published studies of audit and feedback in dementia rarely described more than one cycle. None of the included studies had a comparison group: 12 were before and after designs and one was an interrupted time series without a comparison group. The median absolute improvement was greater than in studies beyond dementia which have used stronger designs with fewer risks of bias. Included studies demonstrated large variation in the effectiveness of audit and feedback.
Conclusions: Whilst methodological and reporting limitations in the included studies hinder the ability to draw strong conclusions on the effectiveness of audit and feedback in dementia care, the large interquartile range indicates further work is needed to understand the factors which affect the effectiveness of this much-used intervention.

What is already known about the topic?

- Audit and feedback is a much-used intervention to implement evidence-based care.
- Audit and feedback varies in the extent to which it changes practice.
- There have been calls for further research to explore the causes of variation in the effectiveness of audit and feedback

What this paper adds

- This review found no evidence that variations in the source, delivery or frequency of audit feedback affected the subsequent care practices in the context of dementia.
- Studies with at least two data collection cycles, which detail the ingredients of the intervention and have a comparator group are

needed to better investigate and understand the factors explaining the variation in the effectiveness of audit and feedback.

1. Background

Evidence-based care for people with dementia is a priority for patients, carers and clinicians (Lind, 2014), and a policy priority (Department of Health, 2012; AHRQ Strategic Plan, 2014). Yet, people with dementia do not always receive evidence-based care. Data from the U.S. suggests gaps in the delivery of evidence-based care, for example, only 31% of people with dementia met the quality indicator of having been screened for depression during the initial evaluation (Arora et al., 2007). In England and Wales, the national audit of dementia found that whilst 97% of organisations had a process in place to undertake a mental state examination, this was done in only 50% of

records audited (Royal College of Psychiatrists, 2013). An earlier audit of hospitals in England and Wales (Souza et al., 2014) found that only 26% of patients with dementia had a standardised assessment of functioning. The challenge of providing evidence-based care for people with dementia is not limited to hospitals, for example, they are often prescribed anti-psychotic drugs inappropriately in care homes (All-Party Parliamentary Group on Dementia, 2008; Kheun, 2013).

Audit and feedback, also known as clinical, nursing or medical audit, is a practice change intervention to increase the receipt of evidence-based care (Souza et al., 2014). Audit and feedback is widely advocated as a way to increase the use of evidence and to provide data to assess assurance of care quality. It is a requirement of professional registration (e.g. General Medical Council, 2012), of regulatory arrangements (e.g. Care Quality Commission, 2010) and has an important role in nursing care (e.g. Christina et al., 2016). Audit and feedback involves comparing current care against an evidence-based standard, and giving feedback to staff on whether that current care meets those standards. Steps in the audit process have been described as planning, standard setting, measuring performance, providing feedback, implementing change and re-auditing (Benjamin, 2008).

A recent Cochrane review of audit and feedback (Ivers et al., 2012) in different clinical domains found that while audit and feedback led to a relatively small 4.3% median absolute improvement in practice (for example, prescribing or undertaking assessments), it also had the potential to result in larger effects, with an interquartile range from 0.5% to 16%. Different theories have been applied to explain how audit and feedback might change behaviour. For example, control theory (Carver and Scheier, 1982) describes how feedback influences the perception of a gap between current and intended behaviour. Persons will seek to close this gap unless they believe it to be too large or they lack the skills or motivation to close it. Under this theory, clear performance feedback and action plans to meet the goal would increase the effectiveness of audit and feedback. Feedback intervention theory (Kluger and DeNisi, 1996) describes that attention is limited and that feedback directs attention to an action, which impacts upon performance of that action. Under this theory, characteristics of the feedback, the task, the actors and the goals determine the extent to which feedback affects performance. Understanding the differential effectiveness of audit could develop understanding of how it works (Ivers et al., 2012) and provide nurses with better information about undertaking effective audit (e.g. Christina et al., 2016). The Cochrane review (above) identified five potential determinants of the effect of audit and feedback as demonstrated within randomised controlled trials. A further search for determinants of the effectiveness of audit and feedback has drawn upon experience and expert interviews (Brehaut et al., 2016). The current review extends the previous searches by exploring influences described within studies that use a range of designs. The specific purpose was to inform work to enhance audit in dementia care and related research, including hypothesis generation for future effectiveness trials.

Specifically, this review investigated:

(1) Whether audit and feedback is effective for improving health professionals' care of people with dementia?
(2) Whether the content and delivery of audit and feedback affect its effectiveness in the context of health professionals' care for people with dementia?

2. Method

A systematic review with adherence to established principles (e.g. Khan et al., 2003) for such reviews was undertaken.

2.1. Protocol and registration

The review protocol provides the details for the searches and is available on Prospero at: http://www.crd.york.ac.uk/PROSPERO/ display_record.asp?ID=CRD42015023688

2.2. Eligibility criteria

The population under study was health professionals providing care to people with dementia and/or the informal carers of people with dementia. The intervention was audit and feedback. The outcome was change in target behaviour (for example, a change in the use of nutrition assessments for people with dementia pre- and post- audit and feedback). To be included, studies were therefore required to have at least two cycles of data. Studies with or without a comparison group were included.

2.3. Information sources

The Cochrane Central Register of Controlled Trials, Prospero, Medline (1946–December week 1 2016), PsycInfo (1967 to January 2017), Cinahl (1982 to January 2017), HMIC (1979 to January 2017), Embase (1974–2017 week 1) databases and the Science Citation Index and Social Science Citation Index were searched.

2.4. Search

The search combined terms relating to dementia, health personnel and audit and feedback. The search strategy was adapted from previous reviews (Ivers et al., 2012; Forbes et al., 2014). An example search is presented in Supplementary Text S1. Peer-reviewed journals, reports, book chapters, theses and conference abstracts which met the following criteria were included.

Reference lists from all included studies and a previous review of audit and feedback (Ivers et al., 2012) were hand-searched and citations of included studies sought using the Science and Social Sciences Citation Indexes and Google Scholar. The resulting references and citations were assessed against the inclusion/exclusion criteria. The search was re-run in January 2017 prior to final reporting in order to identify recently published studies. The reference data was managed using EndNote.

2.5. Study selection

Two researchers (MS, NK) undertook an iterative process involving three cycles of independently screening, assessing agreement, clarifying the criteria and then re-screening against inclusion/exclusion criteria. 131 papers, including 83 abstracts only, were reviewed in this way in order to enable learning and refinement to the approach. Once a consistent approach between the two reviewers was developed, the remainder (4517 papers) were screened by one person (MS). Studies about which the screener was uncertain were discussed by MS and NK. A consistent approach to dual screening was assessed using Kappa coefficient. Titles were screened and either rejected or progressed to abstract screen. The screened abstracts were either rejected or progressed to full-text screen. The full-text screen identified the studies to be included.

2.6. Data collection process

MS and JM independently assessed study design of included studies (Hartling et al., 2011).

Data extraction was undertaken using a pre-determined and piloted tool (Supplementary text S2). Within this, the target behaviours within the studies were categorised as relating to the outcome of care (for example, whether patients were restrained), the process of care (for example, whether there was a care plan for the use of restraint or whether it was provided as per instructions) and structure (for example, training) (Benjamin, 2008). The Template for Intervention Description and Replication (TIDieR) (Hoffmann et al., 2014) was used to

categorise the reporting of the intervention, such as how it was delivered and by whom.

2.7. Data items

Health professionals included those staff with and without a clinical qualification providing care in a healthcare setting, for example, a hospital support worker.

The intervention, audit and feedback, was defined as a combination of (i) collecting/using data about current care to compare the current care against an evidence-based standard and (ii) giving feedback to staff about the match between the current care and evidence-based care. Studies were excluded if audit and feedback was not core to the intervention, for example, where the intervention was described as training with performance data gathered before and after but where this data were used to evaluate the training rather than to provide feedback to the participants.

The outcome was absolute improvement in targeted behaviour. For compliance with a dichotomous standard (for example, the care plans that met an expected standard), absolute improvement in individual target behaviours were calculated as the post-intervention compliance minus pre-intervention compliance. One uncontrolled study presented continuous variable data, this was described narratively. As no studies included control groups, relative effect was not calculated. No adjustment was made for baseline performance.

2.8. Risk of bias in individual studies

The Cochrane risk of bias tool (Higgins et al., 2011) was piloted and amended: MS and JM assessed the studies, discussed the outcomes, refined the Cochrane tool, and re-assessed. Following the re-assessment, areas of disagreement were discussed, and agreement sought. A score of zero was given for a high or unclear rating; a score of one for a low rating. The refinement added a 'medium' risk of bias (0.5 score) so as to differentiate between studies, for example, those using a whole population and those which selected non-randomly were scored as medium and high for selection bias respectively. Scores were summed to give a total out of six for each study.

2.9. Summary measures

The data were described statistically through the calculation of the median absolute improvement and interquartile range.

2.10. Synthesis of results

The review included diverse study designs and presented the study findings alongside their risk of bias and details of the intervention design. The data was then synthesised graphically, including using a harvest plot to explore the heterogeneity between studies (Ogilvie et al., 2008) and present evidence for different hypotheses. The same risk of bias tool was used for all designs in order to support synthesis. The harvest plots were adapted from previous use due to lack of information about statistical differences between subgroups. As a result, the presence or absence of potential influences is located on adjacent plots. For example, the graph of studies involving peer feedback is next to the one showing non-peer feedback. The harvest plots were further adapted to explore the degree of absolute improvement. Studies that provided results as a proportion were included in the graphical synthesis, pooled by type of outcome; where this was not possible, studies were explored narratively. The null hypothesis was that the content and delivery of the audit and feedback intervention leads to no change in the target behaviours under study. The decision to accept or reject the null hypothesis was based on a visual assessment of difference between paired graphs, for example, the plot for feedback being provided once and the plot for repeated feedback.
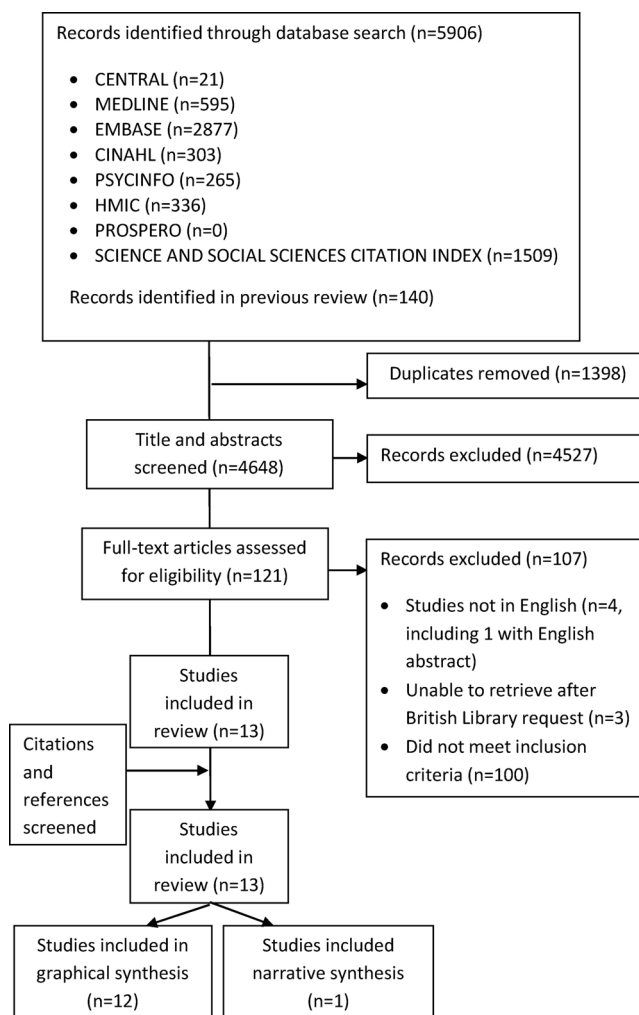

Fig. 1. PRISMA Flow diagram.

3. Results

4508 potential titles were identified, of these thirteen were ultimately included in the review (Fig. 1).

Within the included studies, none had a comparison group: twelve were before and after designs, and one was an interrupted time series without a comparison group. Further details on the included studies is provided in Table 1.

Supplementary table S3 describes the assessment of bias for each included study.

There was a high risk of selection bias for three studies, and a medium risk for seven studies, the risk was unclear for the other three studies. There was a high risk of performance bias for all studies. There was a high risk of detection bias for eleven studies, low risk for one and unclear for one study. There was a medium risk of attrition bias for one study, the risk was low for eight and unclear for a further four studies. There was a high risk of reporting bias for three studies, there was a medium risk for eight studies and the risk was low for one and unclear for a further study. High risk of other biases were identified for four studies.

Two studies (Altus et al., 2002; Rooney, 2014) had a risk of bias score of 2.5 out of 6, indicating that they have fewer risks of bias than the other studies but still had multiple sources of bias. Five studies (Baker and Rogers, 2005; Barton et al., 2006; Dawson, 2014; Knox, 2007; Timmins, 2008) had a risk of bias score of 2, one study (Wilson et al., 2015) had a risk of bias score of 1.5, three studies (Dodd, 2013; Georg, 2006; Reynolds et al., 2006) had a risk of bias score of 0.5 and

Table 1
A table to describe studies meeting the inclusion criteria.

First Author	Study design	Process measures	Baseline/Follow-up (% unless indicated)	Setting as described	Number of centres	Method for selecting staff participants	Number of patients	Number of health personnel
Altus et al., 2002	Interrupted time series	Activity prompts / Praise / Contacts	3.7/26.8 per hour / 1.3/3.2 per hour / 28.3/51.5 per hour	Locked dementia care unit in U.S. at assisted-living facility	1	Purposive from team	6	1
Baker and Rogers, 2005	Before and after	Driving discussed / Discussion followed-up	23.0/95.0 / 33.0/100	Day Hospital in England	1	Whole team	82	Unclear
Barton et al., 2006	Before and after	Etiology of dementia / Management plan / Other provider plan	52.0/91.0 / 35.0/90.0 / 22.0/62.0	Nursing home care unit in U.S. medical centre	1	Whole team	45	Unclear
Dawson, 2014	Before and after	No medication record errors	96.5/99.5	2 dementia assessment units in Northern Irish hospitals	2	Whole team	Unclear	Unclear
Dodd, 2013	Before and after	Confusion assessment	5.6/43.2	Emergency admissions to general surgery in English hospital	Unclear	Unclear	80	Unclear
Georg, 2006	Before and after	Denture care [a]	29.6/46.4	Residential Aged Care Facilities (high and low care) in Australia	4	Unclear	79	121
Knox, 2007	Before and after	Compliant restraint plan / Documented assessment / Documented rationale / Applied as instructions	5.6/22.5 / 52.0/68.6 / 4.5/42.6 / 31.7/84.3	4 high care facilities in Australia	3	Whole team	174	162
Mills et al., 2014	Before and after	Abbreviated mental test score / Onset recorded / Review of medication / Signed and dated / National target met	28.0/46.0 / 10.0/22.7 / 10.0/25.0 / 54.0/76.0 / 12.0/33.0	3 care of the older person wards in English Hospital	3	Purposive	137	Unclear
Morrison, 2009	Before and after	Medication indication record / Documented review	50.0/100 / 64.0/100	3 nursing homes registered within one general practice in Scotland	1	Unclear	37	Unclear
Reynolds et al., 2006	Before and after	Effective interventions	97.2/81.8	Continuing care ward in English hospital	1	Purposive	8	Unclear
Rooney, 2014	Before and after	Standardised tool used / Informed management plan / Non-pharmacological plan / Re-assessment / Response documented / Appropriate prescribing	100/100 / 83.0/100 / 0/100 / 83.0/100 / 83.0/100 / 21.0/100	18-bed Secure dementia unit in Australia	1	Whole team	18	29
Timmins, 2008	Before and after	Compliant restraint plan / Restraint need assessed / Rationale recorded / Applied as instructions	0/100 / 84.8/94.0 / 0/100 / 87.0/88.0	Secure residential aged high care facility in Australia	1	Whole team	Unclear	Unclear

Table 1 (continued)

First Author	Study design	Process measures	Baseline/Follow-up (% unless indicated)	Setting as described	Number of centres	Number of patients	Method for selecting staff participants	Number of health personnel
Wilson et al., 2015	Before and after	Prescription writing assessed against 28 standards	Median improvement = 7.0	Continuing care inpatient unit in Ireland	1	82	Whole team	Unclear

Note: Mean figures presented in this summary table for studies where individual site data provided. [a] = Additional standard excluded due to lack of feedback.

two studies (Reynolds et al., 2006; Rooney, 2014) had a risk of bias score of 0, meaning that for each assessed source of bias it was either present or unclear.

All studies had a risk of at least three sources of bias (Supplement S3), with performance (n = 13) and detection (n = 11) biases being most common.

The included studies described audit and feedback in a range of settings including nursing care homes, day hospitals, and in-patient settings. The studies were from the US, Australia, the UK (England, Northern Ireland and Scotland), and Ireland. Nine of the studies involved nursing staff, with a further three studies describing undefined team involvement. All studies sought change in structures and processes; only one (Timmins, 2008) also sought change in outcome (receipt of restraint). Overall, across the 13 included studies, there was insufficient evidence to reject the null hypothesis that the content and delivery of the audit and feedback intervention leads to no change in the target behaviours under study. Fig. 2 presents the analysis from the 12 studies that provided results as a proportion. Additional plots exploring interactional effects of verbal, written and verbal, once only and repeated feedback were produced without demonstrating differences.

The one study that could not be included in the graph was an uncontrolled interrupted time series without a comparison group (Petticrew et al., 2013); this study described increases in the hourly number of activity prompts (3.7 to 26.8) and praise (1.3 to 3.2) as a result of audit and feedback by the team's activity director. Inappropriate engagement of residents fell from 8.5 to 1.1 and resident engagement in activity increased from 1.4 to 2.3 activities per resident per hour.

Within the 12 studies with proportional data, the median absolute improvement was 22%, inter-quartile range 2.75–52.9%. For the measures of process, the median absolute improvement was 17%, inter-quartile range 0.5–50%. For the seven measures of structure, the median absolute improvement was 72%, inter-quartile range 36.1–82.5%. For the one study (Timmins, 2008) that sought improvement in outcome the absolute improvement was 22%. Twelve studies showed improvement over time, with one study (Reynolds et al., 2006) finding performance deteriorated.

Across the 13 included studies, 67% of the information sought against the TIDieR framework could be identified (Table 2). Nine studies (Baker and Rogers, 2005; Barton et al., 2006; Dawson, 2014; Dodd, 2013; Mills et al., 2014; Morrison, 2009; Reynolds et al., 2006; Timmins, 2008; Wilson et al., 2015) did not present the number of health personnel participants and two (Dawson, 2014; Timmins, 2008) did not describe the number of patients whose care was assessed. Only one of the studies (Reynolds et al., 2006) referenced theory, namely 'learning theory', although four further studies (Georg, 2006; Knox, 2007; Rooney, 2014; Timmins, 2008) referenced application of the Joanna Briggs Getting Research into Practice (GRIP) framework. The GRIP studies demonstrated a 27% median absolute (IQR = 9.45–81.6) improvement, compared with 22% (IQR = 0–50) in the non-GRIP studies. The one study that reported using learning theory showed a reduction in performance. The other substantial gaps in reporting related to planned and actual fidelity (described in none of the included studies), tailoring (described in only 2 studies (Altus et al., 2002; Georg, 2006)) and modifications (described in 6 studies (Altus et al., 2002; Dawson, 2014; Georg, 2006; Knox, 2007; Reynolds et al., 2006; Timmins, 2008)). Whilst materials, procedures, who was involved, and the study location were more evident, they often lacked the detail necessary to enable further hypotheses to be explored graphically.

4. Discussion

We searched literature to identify whether audit and feedback is effective for improving health professionals' care of people with dementia, and whether the content and delivery of audit and feedback affects its effectiveness in this context. The median absolute
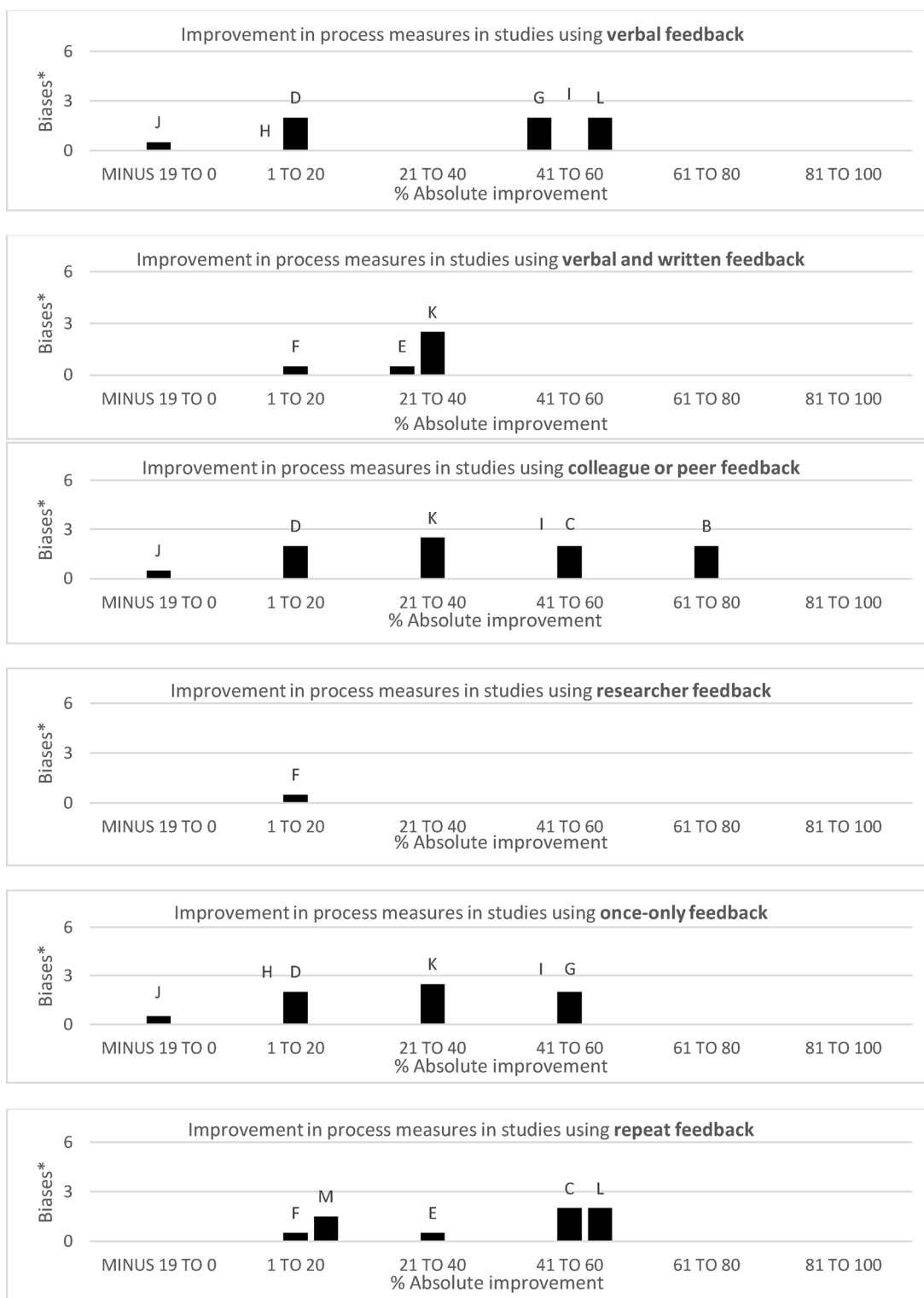
* Each bar represents an included study. Increased bar height indicates lower risk of bias (Max score=6)

B. Baker et al [21]	E. Dodd [24]	H. Mills et al [27]	K. Rooney [30]
C. Barton [22]	F. Georg [25]	I. Morrison [28]	L. Timmins [31]
D. Dawson [23]	G. Knox [26]	J. Reynolds et al [29]	M. Wilson et al [32]

Fig. 2. Harvest plot to describe affects upon improvement.

improvement was 22% (interquartile range 2.75–52.9%).

The absolute improvement from audit and feedback in the present review is much larger than that identified in previous systematic reviews (Ivers et al., 2012; Jamtvedt et al., 2003; Thomson Obrien et al., 2003), the most recent of which found 4.3% absolute improvement (interquartile range 0.5–16%). This may be because the previous

Table 2
Selected information from the data extraction against TIDieR criteria.

First Author	Theory	Materials	Procedure	Who gave feedback	How: Verbal/ Written Feedback	How much: Once/ Repeated Feedback	Tailoring	Modification
Altus et al., 2002	No	Described	Described	Colleague	Unclear	Repeated	One-to-one feedback	Change to method
Baker and Rogers, 2005	No	Not described	Described	Colleague	Unclear	Unclear	Not described	Not described
Barton et al., 2006	No	Described	Described	Peer/Colleague	Unclear	Repeated	Not described	Not described
Dawson, 2014	No	Described	Described	Peer/Colleague	Verbal	Once	Not described	Change to method
Dodd, 2013	No	Described	Described	Unclear	Both	Repeated	Not described	Not described
Georg, 2006	GRIP[i]	Described	Described	Researcher	Both	Repeated	One-to-one feedback	Change to method
Knox, 2007	GRIP[i]	Described	Described	Unclear	Verbal	Once	Not described	Changed differently between sites
Mills et al., 2014	No	Described	Described	Unclear	Verbal	Once	Not described	Not described
Morrison, 2009	No	Described	Described	Colleague	Verbal	Once	Not described	Not described
Reynolds et al., 2006	Learning theory	Described	Described	Colleague	Verbal	Repeated	Not described	Change to method
Rooney, 2014	GRIP[i]	Described	Described	Colleague	Both	Once	Not described	Not described
Timmins, 2008	GRIP[i]	Described	Described	Colleague	Verbal	Repeated	Not described	Change to feedback
Wilson et al., 2015	No	Described	Described	Unclear	Unclear	Repeated	Not described	Not described

Note: All studies had a brief name. Details about setting are provided in Table 1. Reynolds et al., 2006 noted feedback was given immediately, no other studies described timeliness. No studies described planned or actual fidelity. Materials and procedures refer to information about stages within the audit cycle from which data for graphical comparison were sought. It is not possible to include the full description captured within data collection here. Peer refers to someone from same profession but not necessarily same hierarchical level. Colleague refers to someone who works with the recipient. Repeated means more than once.
[i]GRIP = Joanna Briggs' Getting Research into Practice framework.

reviews applied a weighted median and included high-quality randomised controlled trials while the current review included weaker designs in order to explore potential influences on effectiveness. An important finding from our study is the lack of studies of audit and feedback in dementia which demonstrate a low risk of bias. All studies had at least three sources of bias; this undermines the ability to draw strong conclusions about the extent of and variation in the effectiveness of audit and feedback.

Our review identified studies of audit being used to drive improvement across dementia care domains in different care systems. This provides further evidence that understanding what makes audit and feedback more effective has the potential to improve significantly the care received by people with dementia. Simultaneously, reducing the use of less effective audit and feedback approaches will save resources. The review also found that few studies described two cycles of audit. Whilst studies with only a single cycle may provide valuable descriptive data, they prevent assessment of improvement. Few studies described giving feedback to those providing care. This represents either a gap in reporting or a missed opportunity to explore the role feedback has in supporting improvement. Only one study included the measures needed to determine whether the benefits of audit and feedback impact upon patient outcome.

The audit and feedback study method reflects the assumptions of those involved in the design. Understanding the 'theory incarnate' (Pawson and Tilley, 1997) assists the reader to understand the ingredients of the intervention. Applying existing theories builds upon existing understanding and helps develop and test these theories (Moore et al., 2014). Within the included studies, the one study to report using learning theory (Timmins, 2008) demonstrated deterioration in performance. There was some very weak evidence of the use of the Getting Research Into Practice framework (e.g. Rooney, 2014) to increase effectiveness, but overall the identification of specific factors explaining variation in the effectiveness of audit and feedback was hampered by gaps in reporting. There remains a need to ensure authors of primary studies explicitly describe interventions in terms of who did what and how (Hoffmann et al., 2014).

The use of harvest plots enabled an exploration of potential determinants of the effectiveness of audit and feedback, although none were identified from the included papers. A more fine-grained review of the papers studying multiple sites found potential contextual determinants of effectiveness. Within one nurse-led audit of restraint in Australian residential high-care facilities using the Getting Research Into Practice framework (Knox, 2007), different levels of compliance were achieved across three different sites, for example, unlike sites A and B, site C gained 100% compliance at the second cycle in three standards. A detailed examination of the differences between sites in this paper (Knox, 2007), highlights contextual determinants of effectiveness worthy of further exploration: Sites A and B both had approximately 70 residents, whilst site C was, "in the process of closing down" (p103) and the population in site C changed to have fewer people with dementia or impaired mobility. As part of the change process, stakeholders developed a "new policy, procedure, authorisation form, restraint release form, information sheet for residents and families, and a self-directed learning package for direct care staff" (p105). This was piloted in site C, which if time to implement the recommendations is an important influence, might explain them reaching full compliance for the documentation of assessment and rationale and the application of restraints as per instructions. However, site C also had the lowest staff education figure in cycle 2, which suggests that whilst a specific recommendation might make change more likely (e.g. Brehaut et al., 2016) other factors may moderate and/or mediate this effect. The organisational policy across sites changed between cycles such that bed rails were classed as restraints. The paper describes how this had a disproportionate influence upon site A, but does not expand upon this. The paper (Knox, 2007) also describes staff, resident and family anxieties about the changes being made as a result of the audit, although it does not note whether these differed between sites. This is regrettable as there may be theoretically important implications, such as a link to local goals and priorities (e.g. Locke and Latham, 1990).

Two other studies (Dawson, 2014; Georg, 2006) describe performance at more than one site. Georg et al. (Georg, 2006) provide little detail to explain differences between sites, although they note that resident choice and a short gap between audit cycles may have reduced the opportunity for improvement. Dawson et al. (Dawson, 2014) describe differential improvement between in-patient sites, noting that one site did not improve between cycle 1 and 2, but did after a third cycle. The improvement at the third cycle was attributed to a change in ward management and on-site pharmacy presence. They also note that, "nursing staff were recruited from ward A to ward B to assist with data collection in audits 2 and 3. This appeared to improve attitudes towards

the audit, and further highlighted the problem to nursing staff" (p22). It is possible therefore that initial perceptions towards the audit process may affect effectiveness, and/or this may provide further evidence for the role of repeated feedback. The fourth study (Mills et al., 2014) using multiple sites provided cumulated data that prevents exploration of difference.

In summary, we sought to build upon the findings of the Cochrane review by not excluding based upon study design. Our findings were undermined by bias within included studies and gaps in reporting. However, we were able to draw tentative findings based upon those that used the Getting Research Into Practice framework and from the detailed review of papers that studied multiple sites: Future research should consider potential determinants of effectiveness within the audit planning stage, as well as the structure and delivery of the feedback and action planning. For example, whether feedback is more effective when it involves a wider discussion about barriers to change. In addition, a process evaluation exploring staff attitudes towards the audit and the extent of staff agreement with changes may extend our understanding of the determinants of the effect of audit and feedback.

5. Limitations

Like all reviews, studies included in this review are at risk of being affected by publication bias. Searches of multiple databases were undertaken to reduce the risk of missing studies, however it remains possible that hand-searching may have identified further studies screened out by the search strategy. The final iteration of dual screening demonstrated excellent agreement (Kappa 0.887), however only a small sample were dual screened and it is possible that screening changed over the remaining papers resulting in papers being inappropriately excluded. Due to exclusion by language (Bourgeois, 2008; Koskas et al., 2009; Pancrazi, 2005; Prado Somalo et al., 2006) and unavailability at the British Library, three abstracts and four full-text papers were not screened. Contacting authors may have reduced gaps in reporting but at the cost of potential response bias. In the absence of a gold-standard approach for assessing study design, the authors used a tool (Hartling et al., 2011) which may have resulted in misclassification. The high risk of bias within the included studies mean that the effect size identified must be viewed with caution. In order to draw conclusions it was necessary to group different target behaviours, such as documenting the use of restraint and of a particular assessment tool being grouped as process measures. Defining restraint in use as an outcome and being prescribed anti-psychotics (but without information about whether it was given) as a process may not be meaningful to patients. No assessment was made in relation to the evidence for the target behaviour nor the relationship between process measures and patient outcomes.

Audit and feedback may have a differential effect on different behaviours and pooling effects across behaviours may hide true effects. The gaps in reporting numbers of participants prevented calculation of weighted median absolute improvement. Multi-facetted interventions (e.g. Rolland et al., 2016) where audit and feedback was not core were excluded, but it is possible that further evidence about the effectiveness of audit and feedback could have been identified from these excluded papers. The use of adapted harvest plots in the present review was a novel solution for describing and accounting for sources of biases, however gaps in reporting meant that the plots lack information about the effect of further variables and interactions, for example, exploring repeated verbal feedback from a colleague against that from a researcher.

6. Conclusion

Overall, there is a lack of robust evidence about factors that affect the effectiveness of audit and feedback in dementia care. Given gaps in the current delivery of care, the frequency of audit and feedback use and its variable effectiveness, there is an opportunity to improve care

by making potentially simple enhancements to how current audit and feedback is undertaken. To achieve this, evidence from head-to-head intervention trials using randomised study designs, reported more fully than is currently the case within dementia care and supported by a process evaluation is needed.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.ijnurstu.2017.10.013.

References

AHRQ Strategic Plan, Content last reviewed October 2014. Agency for Healthcare Research and Quality, Rockville, MD. Available at: http://www.ahrq.gov/cpi/about/mission/strategic-plan/strategic-plan.html Accessed 6.1.2017.

All-Party Parliamentary Group on Dementia, 2008. Always a Last Resort: Inquiry into the Prescription of Antipsychotic Drugs to People with Dementia Living in Care Homes. Alzheimer's Society, London.

Altus, D.E., Engelman, K.K., Mathews, R.M., 2002. Finding a practical method to increase engagement of residents on a dementia care unit. Am. J. Alzheimer's Dis. Dement. 17 (4), 245–248.

Arora, V.M., Johnson, M., Olson, J., et al., 2007. Using assessing care of vulnerable elders quality indicators to measure quality of hospital care for vulnerable elders. J. Am. Geriatr. Soc. 55, 1705–1711.

Baker, P., Rogers, T., 2005. Audit of clinicians' approach to patients' driving status in dementia day hospital setting. Psychiatric Bull. 29 (10), 372–373.

Barton, C., Miller, B., Yaffe, K., 2006. Improved evaluation and management of cognitive impairment: results of a comprehensive intervention in long.-term care. J. Am. Med. Dir. Assoc. 7 (2), 84–89.

Benjamin, A., 2008. Audit: how to do it in practice. BMJ 336, 1241–1245.

Bourgeois, F., 2008. To improve the daily care of Alzheimer patients in the general hospital. Rev. Infirm. 143 (September), 23–24.

Brehaut, J., Colquhoun, H., Eva, K., et al., 2016. Practice feedback interventions: 15 suggestions for optimizing effectiveness. Ann. Intern. Med. 164, 435–441.

Care Quality Commission, 2010. Essential Standards. CQC, London.

Carver, C.S., Scheier, M.F., 1982. Control theory: a useful conceptual framework for personality-social, clinical, and health psychology. Psychol. Bull. 92 (July (1)),

111–135.

Christina, V., Baldwin, K., Biron, A., et al., 2016. Factors influencing the effectiveness of audit and feedback: nurses' perceptions. J. Nurs. Manag. 24, 1080–1087.

Dawson, E., 2014. Reducing medicines administration errors on dementia units. Prog. Neurol. Psychiatry 18 (10), 20–23.

Department of Health, 2012. Prime Minister's Challenge on Dementia: Delivering Major Improvements in Dementia Care and Researchby 2015. Department of Health, London.

Dodd, C., 2013. An audit investigating cognitive assessment of emergency general surgical admissions in elderly patients in Sheffield, UK. Int. J. Surg. 11 (8), 649.

Forbes, D., Blake, C.M., Thiessen, E.J., et al., 2014. Light therapy for improving cognition, activities of daily living, sleep, challenging behaviour, and psychiatric disturbances in dementia (Review). Cochrane Database Syst. Rev(Accessed 6.1.2017) Available at: http://onlinelibrary.wiley.com/doi/10.1002/14651858.CD003946.pub4/full.

General Medical Council, 2012. Supporting Information for Appraisal and Revalidation. General Medical Council, Manchester.

Georg, D., 2006. Improving the oral health of older adults with dementia/cognitive impairment living in a residential aged care facility. Int. J. Evid. Based Healthcare 4 (1), 54–61.

Hartling, L., Bond, K., Harvey, K., Santaguida, L., et al., 2011. Testing a tool for the classification of study designs in systematic reviews of interventions and exposures showed moderate reliability and low accuracy. J. Clin. Epidemiol. 64, 861–871.

Higgins, J.P.T., Altman, D.G., Gotzsche, P.C., et al., 2011. Cochrane bias methods group. cochrane statistical methods group. the cochrane collaboration's tool for assessing risk of bias in randomised trials. BMJ 343.

Hoffmann, T.C., Glasziou, P.P., Boutron, I., et al., 2014. Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. BMJ 2014.

Ivers, N., Jamtvedt, G., Flottorp, S., et al., 2012. Audit and feedback: effects on professional practice and healthcare outcomes'. Cochr. Database Syst. Rev. 6http://dx.doi.org/10.1002/14651858.CD000259.pub3. Art. No.: CD000259.

Jamtvedt, G., Young, J.M., Kristoffersen, D.T., et al., 2003. Audit and feedback: effects on professional practice and health care outcomes. Cochr. Database Syst. Rev. 2.

Khan, K.S., Kunz, R., Kleijnen, J., et al., 2003. Five steps to conducting a systematic review. J. R. Soc. Med. 96 (3), 118–121.

Kheun, B.M., 2013. Efforts stall to curb nursing home antipsychotic use. J. Am. Med. Assoc. 310 (11), 1109–1110.

Kluger, A.N., DeNisi, A., 1996. The effects of feedback interventions on performance: a historical review, a meta-analysis, and a preliminary feedback intervention theory. Psychol. Bull. 119, 254–284.

Knox, J., 2007. Reducing physical restraint use in residential aged care: implementation of an evidence-based approach to improve practice. Int. J. Evid. Based Healthcare 5 (1), 102–107.

Koskas, P., Verstichel, P., Fathi, Z., Drunat, O., Wolmark, Y., 2009. Audit cible des indications des électroencéphalogrammes dans un hôpital gériatrique sur une période de 16 mois. Presse Med. 38, 1860–1864.

James Lind Alliance Priority Setting Partnerships, 2014. Available at http://www.jla.nihr.ac.uk/priority-setting-partnerships/dementia/top-10-priorities/ Accessed 6.1.2017.

Locke, E.A., Latham, G.P., 1990. A Theory of Goal Setting and Task Performance. Prentice Hall, Englewood Cliffs, NJ.

Mills, J., Minhas, J., Robotham, S.L., 2014. An assessment of the dementia CQUIN − an audit of improving compliance. Dement. Int. J. Social Res. Pract. 13 (5), 697–703.

Moore, G., Audrey, S., Barker, M., et al., 2014. Process Evaluation of Complex Interventions: UK Medical Research Council (MRC) Guidance. (Accessed 20 6.17) Available at: https://www.mrc.ac.uk/documents/pdf/mrc-phsrn-process-evaluation-guidance-final/.

Morrison, A., 2009. Antipsychotic prescribing in nursing homes: an audit report. Qual. Primary Care 17 (5), 359–362.

Ogilvie, D., Fayter, D., Petticrew, M., et al., 2008. The harvest plot: a method for synthesising evidence about the differential effects of interventions'. BMC Med. Res. Methodol. 8, 8.

Pancrazi, M.P., 2005. Benchmarking Alzheimer's Dementia in Institutions, vol. 52. pp. 42–45 March–April.

Pawson, R., Tilley, N., 1997. Realistic Evaluation. Sage, London.

Petticrew, M., Rehfuessb, E., Noyesc, J., et al., 2013. Synthesizing evidence on complex interventions: how meta-analytical, qualitative, and mixed-method approaches can contribute. J. Clin. Epidemiol. 66, 1230–1243.

Prado Somalo, M., Caceres Camblor, A., Hernandez Zorzano, N., Mendoza Quitanar, R., Ozaeta Alonso, N., Saenz Abad, I., Lopez Sancho, V., 2006. Clinical audit regarding the administration of drugs in a residence for elderly people with Alzheimer's disease and other neurodegenerative dementias. Gerokomos 17 (1), 32–36.

Reynolds, R., King, M., Herzberg, J., 2006. Can a clinical audit improve communication? J. Dement. Care 14 (5), 17–18.

Rolland, Y., Mathieu, C., Piau, C., et al., 2016. Improving the quality of care of long-stay nursing home residents in France. J. Am. Geriatr. Soc. 64, 193–199.

Rooney, C., 2014. Management of physical aggression in the dementia resident in an 18 bed secure dementia unit: a best practice implementation project. JBI Database Syst. Rev. Implement. Rep. 12 (8), 394–410.

Royal College of Psychiatrists, 2013. Report of the National Audit of Dementia: Care in General Hospitals. Healthcare Quality Improvement Partnership, London.

Souza, R., Gandesha, A., Hood, C., et al., 2014. Quality of care for people with dementia in general hospitals: national cross-sectional audit of patient assessment. Clin. Med. 14 (5), 490–494.

Thomson Obrien, M.A., Oxman, A.D., Davis, D.A., et al., 2003. Audit and feedback: effects on professional practice and health care outcomes. Cochrane Database Syst. Rev. 1.

Timmins, J., 2008. Compliance with best practice: implementing the best available evidence in the use of physical restraint in residential aged care. Int. J. Evid. Based Healthcare 6 (3), 345–350.

Wilson, L.S., Devitt, P., Hally, O., 2015. Standards of prescription writing in a long-term psychogeriatric unit: a series of clinical audits. Irish J. Psychol. Med. 32 (2), 197–204.