Published in final edited form as:
Cochrane Database Syst Rev. ; (11): CD004538. doi:10.1002/14651858.CD004538.pub2.

Public release of performance data in changing the behaviour of healthcare consumers, professionals or organisations

Nicole ABM Ketelaar[1], **Marjan J Faber**[1], **Signe Flottorp**[2,3], **Liv Helen Rygh**[2], **Katherine HO Deane**[4], and **Martin P Eccles**[5]

[1]Scientific Institute for Quality of Healthcare (IQ healthcare), Radboud University Nijmegen Medical Centre, Nijmegen, Netherlands. [2]Norwegian Knowledge Centre for the Health Services, Oslo, Norway. [3]Department of Public Health and Primary Health Care, University of Bergen, Bergen, Norway. [4]Edith Cavell Building, University of East Anglia, Norwich, UK. [5]Institute of Health and Society, Newcastle University, Newcastle upon Tyne, UK

Abstract

Background—It is becoming increasingly common to release information about the performance of hospitals, health professionals or providers, and healthcare organisations into the public domain. However, we do not know how this information is used and to what extent such reporting leads to quality improvement by changing the behaviour of healthcare consumers, providers and purchasers, or to what extent the performance of professionals and providers can be affected.

Objectives—To determine the effectiveness of the public release of performance data in changing the behaviour of healthcare consumers, professionals and organisations.

Search methods—We searched the Cochrane Central Register of Controlled Trials (CENTRAL), Cochrane Effective Practice and Organisation of Care (EPOC) Trials Register, MEDLINE Ovid (from 1966), EMBASE Ovid (from 1979), CINAHL, PsycINFO Ovid (from 1806) and DARE up to 2011.

Selection criteria—We searched for randomised or quasi-randomised trials, interrupted time series and controlled before-after studies of the effects of publicly releasing data regarding any aspect of the performance of healthcare organisations or individuals. The papers had to report at least one main outcome related to selecting or changing care. Other outcome measures were awareness, attitude, views and knowledge of performance data and costs.

Contact address: Nicole ABM Ketelaar, Scientific Institute for Quality of Healthcare (IQ healthcare), Radboud University Nijmegen Medical Centre, PO Box 9101, Nijmegen, 6500 HB, Netherlands. n.ketelaar@iq.umcn.nl.
Editorial group: Cochrane Effective Practice and Organisation of Care Group.
Publication status and date: New, published in Issue 11, 2011.
Review content assessed as up-to-date: 26 June 2011.
CONTRIBUTIONS OF AUTHORS NK drafted the protocol and amended it in line with comments from MF, SF, LHR, KD and MPE. NK drafted the search strategy. All authors screened the studies, checked the quality, examined studies for eligibility, extracted data and analysed results. NK drafted the text. All authors commented on sequential drafts.
DECLARATIONS OF INTEREST None known.

Europe PMC Funders Author Manuscripts

Data collection and analysis—Two review authors independently screened studies for eligibility and extracted data. For each study, we extracted data about the target groups (healthcare consumers, healthcare providers and healthcare purchasers), performance data, main outcomes (choice of healthcare provider and improvement by means of changes in care) and other outcomes (awareness, attitude, views, knowledge of performance data and costs).

Main results—We included four studies containing more than 35,000 consumers, and 1560 hospitals. Three studies were conducted in the USA and examined consumer behaviour after the public release of performance data. Two studies found no effect of Consumer Assessment of Healthcare Providers and Systems information on health plan choice in a Medicaid population. One interrupted time series study found a small positive effect of the publishing of data on patient volumes for coronary bypass surgery and low-complication outliers for lumbar discectomy, but these effects did not persist longer than two months after each public release. No effects on patient volumes for acute myocardial infarction were found.

One cluster-randomised controlled trial, conducted in Canada, studied improvement changes in care after the public release of performance data for patients with acute myocardial infarction and congestive heart failure. No effects for the composite process-of-care indicators for either condition were found, but there were some improvements in the individual process-of-care indicators. There was an effect on the mortality rates for acute myocardial infarction. More quality improvement activities were initiated in response to the publicly-released report cards. No secondary outcomes were reported.

Authors' conclusions—The small body of evidence available provides no consistent evidence that the public release of performance data changes consumer behaviour or improves care. Evidence that the public release of performance data may have an impact on the behaviour of healthcare professionals or organisations is lacking.

Medical Subject Headings (MeSH)

*Information Dissemination; *Quality Improvement; Canada; Consumer Health Information [*methods]; Evaluation Studies as Topic; Health Maintenance Organizations [standards]; Hospitals [*standards]; Medicaid; Organizational Innovation; Quality Assurance, Health Care [*methods]; Randomized Controlled Trials as Topic; Reproducibility of Results; United States

MeSH check words

Humans

BACKGROUND

It is becoming increasingly common to release information about the performance of healthcare systems into the public domain. In the present era of accountability, cost-effectiveness, quality improvement and demand-driven healthcare systems, policy and decision-makers such as governments, regulators, purchaser and provider organisations, health professionals and consumers of health care are becoming more interested in measuring performance (Smith 2009). The measurements may appear in consumer reports, provider profiles or report cards. It is not always clear who the information users are or what

the release of data is expected to achieve. However, it is often assumed that the information will affect and facilitate the decisions and behaviours of various stakeholders and ultimately result in health system improvements (Marshall 2000; Berwick 2003; Smith 2009).

The stakeholders in this review include healthcare consumers, professionals, providers and purchaser organisations. Accountability relationships connect all the stakeholders. These relationships have two prominent elements, namely the 'provision of information' about performance and the 'sanctions or rewards for the accountable party' (Smith 2009 Chapter 1.1; p.5). The main role of performance measurement is to keep the various agents accountable by enabling stakeholders to make informed decisions (Smith 2009). Various suggested uses of performance measurements are linked to the accountability perspective. Examples of underlying objectives are (1) the use of performance measurements to promote more efficient and demand-driven healthcare (Bentley 1998; Hendriks 2009) and (2) applying the results as a marketing tool (Longo 1997). One user goal is to use the public disclosure of performance measurements to encourage providers to focus on quality problems and to stimulate performance improvement (Lindenauer 2007; Fung 2008; Hendriks 2009). From a healthcare consumer's perspective, the data can encourage patients to preferentially choose high-quality health care, i.e. the best health plan or provider (Hibbard 2009; Kolstad 2009; Werner 2009) or to assess the performance of individual professionals (Marshall 2004; Fotaki 2008). Other proposed user goals for performance measurements have been linked to controlling costs (Berwick 1990; Sirio 1996), regulating the healthcare system (Rosenthal 1998; Schut 2005) and influencing the decisions of healthcare purchasers (Brook 1994; Hibbard 1997; Mukamel 1998).

Consumers must overcome barriers to the use of performance data. Examples of such barriers are the complexity of the performance data (Hibbard 2010), lack of skills to comprehend and use performance data (Hibbard 2001; Magee 2003; O'Meara 2005; Peters 2007; Hibbard 2007) and the way data are presented (Harris-Kojetin 2007; Peters 2007; Fung 2008; Damman 2010). A negative consequence of such barriers might be related to the impact of choice on equity in healthcare. Consumers from poorer backgrounds and with lower educational levels will be less likely to be given a choice, less able to choose and less able to afford travel to a better performing, but more distant, provider (The King's Fund 2010).

Professionals focus on the barriers to accessibility (Harris 2008), the validity of the performance measures themselves (Giuffrida 1999; Kerr 2007), and the validity of implicit or explicit comparisons of performance (Parry 1998; Rixom 2002). There are concerns that failure to adequately adjust for the case mix in the data sets may lead to hospitals or clinicians who treat higher-risk patients being labelled as poor performers, or to providers preferentially selecting lower-risk patients (Werner 2005a; Dranove 2008; Bardach 2009). In healthcare systems where providers charge for their services, the 'better' performing providers may charge more (Mukamel 1998), thereby restricting access to better care. Publicly releasing performance data may have other unintended consequences as well. There is a risk that the release may lead to improved reporting without necessarily improving performance. It has been said that the care tasks that are easiest to measure are often those

Europe PMC Funders Author Manuscripts

least important in a quality improvement context and that other task measurements will be neglected (Loeb 2004).

Thus, the impact of public release of performance data may have various mechanisms. Most commentators seem to consider the most important goal of publishing performance data to be to cause providers to improve their performance. This goal can be achieved in a selection pathway or a change pathway (Berwick 2003). Consumers, patients and purchaser organisations that are in a position to do so can select the best healthcare professionals and organisations. This type of selection will not change the quality of the delivered care by itself, but it can be a stimulus for quality improvement. In a change pathway, healthcare professionals and organisations can improve performance by changing their work procedures or professional culture, and organisations can make structural changes.

Description of the intervention

Public release of performance data is the release of information about the quality of care so that patients and consumers can better decide what health care they wish to select and healthcare professionals and organisations can better decide what to provide, to improve or to purchase. This mechanism excludes the use of auditing and feedback as a tool for improving professional practice and healthcare outcomes. This subject has been reviewed elsewhere (Jamtvedt 2006).

Why it is important to do this review

Some systematic reviews (Marshall 2000; Shekelle 2008; Fung 2008; Faber 2009) have suggested positive effects of publicly releasing performance data. However, none of them focuses on identifying and synthesising only the most robust evidence available; this systematic review will do so.

OBJECTIVES

To estimate the effects of publicly releasing performance data on changing the behaviour of three target groups: healthcare consumers (patients), providers of healthcare (health professionals) and purchasers of healthcare.

METHODS

Criteria for considering studies for this review

Types of studies—

- Randomised controlled trials (RCT), including cluster-randomised controlled trials (ClRCTs)
- Quasi randomised trials (QRT), including cluster quasi-randomised trials (ClQ-RCTs) using methods of allocation such as alternation or allocation by case note number.
- Interrupted time series (ITS) studies with at least three data points before and three data points after the intervention.

Europe PMC Funders Author Manuscripts

- Controlled before-after (CBA) studies, with at least two intervention sites and two control sites that are chosen for similarity of main outcome measures at baseline.

Types of participants—Patients or other healthcare consumers and healthcare providers, including organisations (e.g. hospitals, practices and individual healthcare professionals) without any restriction by type of health-care professional, provider, setting or purchaser.

Types of interventions—We included interventions that contained the following elements.

- Performance data about any aspect of the healthcare organisations or individuals, including process measures (e.g. waiting times), healthcare outcomes (e.g. mortality), structure measures (e.g. presence of waiting rooms), consumer or patient experiences (e.g. Consumer Assessment of Healthcare Providers and System (CAHPS) data) and/or expert or peer-assessed measures (e.g. certification, accreditation and quality ratings given by colleagues) (Harris 2002). The data presented may or may not provide comparisons with similar providers or quality standards and may or may not be adjusted for case mix. Performance data may be prepared and released by any organisation, such as the government, insurers or consumer organisations.
- The release of performance data into the public domain in written or electronic form, with varying degrees of accessibility, such as a report in a publicly accessible library or more active dissemination directly to consumers in newspapers, leaflets, personal mailings, broadcasting media, etc.

The data may be presented numerically, graphically or pictorially.

Comparators: The following comparisons were planned.

1. Public release of performance data compared to control (the control intervention should consist of the usual practice in that setting, which may include other interventions aimed at quality improvement, such as the internal use of the same performance data)
2. Different types of public release of performance data compared to each other.

We excluded studies that did not expose participants to performance data concerning process measures, healthcare outcomes, structure measure, consumer/patient experiences or expert or peer-assessed measures. We also excluded studies that reported only hypothetical choices.

Types of outcome measures

Main outcome measures: We planned to the primary outcome measures according to two important aims of those publicly releasing performance data.

1. Improvement by selection:

- Changes in the healthcare utilisation decisions of consumers (public and patients)

- Changes in the healthcare utilisation decisions of healthcare providers (professionals and organisations)
- Changes in the healthcare utilisation decisions of purchasers

2. Improvement by changes in care:

- Objective measures of provider performance, including those that were made public and others that were not
- Valid measures of staff morale or behaviour ('valid' defined as having the development of the assessment tool reported in a peer-reviewed journal).

Other outcome measures: If a study reported at least one main outcome measure we also collected those concerning awareness, attitude, views, knowledge of performance data in all target groups and cost data. Where possible, we planned to collect data about the extent to which outcome measures varied with participant characteristics. We excluded studies that reported awareness, knowledge, attitude or costs in the absence of objective measures of provider performance or decision behaviour of healthcare consumers, providers or purchasers.

Where possible, we planned to collect data about the extent to which outcome measures varied with participant characteristics.

Search methods for identification of studies

We searched the Cochrane Central Register of Controlled Trials (CENTRAL), MEDLINE Ovid (from 1966), EMBASE Ovid (from 1979), CINAHL, PsycINFO Ovid (from 1806) and DARE up to 2011. For MEDLINE, we used subject headings and the relevant quality of healthcare MeSH terms, such as 'process assessment', 'outcome assessment', 'quality indicators', 'quality assurance' and 'benchmarking'. We also used text words and phrases such as 'performance outcome', 'report card', 'criteria', 'standard', 'disclosure', 'quality information' and 'public information'. We combined these terms with forms of decision-making such as 'choice behaviour', 'patient preferences', 'patient acceptance' and 'consumer satisfaction'. We searched the other databases using the appropriate controlled vocabulary. In addition, we identified potentially relevant studies in the reference lists of key articles. Appendix 1 to Appendix 5 give full details of the search terms.

Data collection and analysis

Selection of studies—We downloaded all titles and abstracts (N = 6839) retrieved in the electronic search to a reference management database. We removed the duplicates, then two review authors independently examined the remaining references. All review authors recorded their assessments of abstracts with points: '0' for exclusion, '1' for doubtful and '2' for inclusion. Two review authors independently rated each abstract, therefore a minimum score of '0' and a maximum score of '4' was possible. Abstracts with a combined score of 0 or 1 were excluded. Studies with a combined score of 3 or 4 were included. Two review authors resolved the fate of studies with a combined score of 2 by discussion. A third review author (MPE) decided any disagreements that remained unresolved. We documented

Europe PMC Funders Author Manuscripts

the reasons for exclusion. We obtained full-text copies of papers taken from references for inclusion. Two authors of our review independently assessed the eligibility of these papers.

Data extraction and management—After the first selection round, relevant studies were retrieved for full-text reading (Figure 1). We distributed these studies to our authors in such way that they did not receive studies for text reading that they already evaluated in the first round. We extracted the data about the study design, patient and provider characteristics, interventions, outcome measures, and healthcare choices to a form specially designed for our review (Appendix 6). We used another form (Appendix 7) for the studies we retrieved for a more detailed evaluation. The two review authors resolved disagreements by discussion if possible. A third review author (MPE or MF) dealt with disagreements that the two review authors could not resolve.

Assessment of risk of bias in the included studies—We assessed the risk of bias on the basis of the Cochrane Collaboration criteria (Higgins 2011): (i) adequate sequence generation, (ii) concealment of allocation, (iii) blinding, (iv) incomplete outcome data, (v) selective reporting and (vi) no risk of bias from other sources. We used three additional criteria that the Cochrane Effective Practice and Organisation of Care (EPOC) Group specifies (EPOC 2009): (vii) baseline characteristic similarity, (viii) reliable primary outcome measures and (ix) adequate protection against contamination. We used these nine standard criteria for ClRCTs, ClQ-RCTs and CBA studies. We used seven criteria for ITS studies: (i) the intervention is independent of other changes, (ii) the shape of the intervention effect is pre-specified, (iii) the intervention is unlikely to affect data collection, (iv) knowledge of the allocated interventions is adequately prevented during the study, (v) the outcome data are incomplete, (vi) reporting is not selective and (vii) there is no risk of bias from other sources. Two review authors independently examined the risk of bias assessment and resolved disagreements by discussion. There were some disagreements about the rating of the criterion as ‘yes’ instead of ‘unclear’ or vice versa. Sometimes the rating was based on a different phrase in the text. A third review author (MPE or MF) dealt with any disagreements that the two review authors could not resolve.

Unit of analysis issues—We noted whether studies randomised patients or healthcare providers. If analysis did not allow for clustering of patients within healthcare providers, we recorded a unit of analysis error, because such analyses tend to overestimate the precision of the treatment effect.

Data synthesis—We report the effect sizes for each outcome for each study. Quantitative synthesis was not possible.

RESULTS

Description of studies

See: Characteristics of included studies; Characteristics of excluded studies.

See tables of ‘Characteristics of included studies’ and ‘Characteristics of excluded studies’.

Europe PMC Funders Author Manuscripts

Results of the search—The searches found 6839 references; we excluded 6786 references because the titles and abstracts did not meet our inclusion criteria. We retrieved the full texts of publications listed in 53 references. Four citations were irretrievable and could not be considered for inclusion in the review. An additional search brought two more references to light. Altogether, we retrieved full text versions of 51 papers. Forty-two of these papers did not fulfil the inclusion criteria. We evaluated nine papers in more detail, and subsequently we excluded another five papers. We reported the reasons for exclusion of these 51 studies (including the four irretrievable citations) in the 'Characteristics of excluded studies' table. Four papers met the inclusion criteria of the review. Figure 1 presents the study flow chart (Moher 1999).

Included studies

Characteristics of setting and patients/consumers: See Characteristics of included studies and Table 1.

We included four studies (Farley 2002a; Farley 2002b; Romano 2004; Tu 2009) comprising more than 35,000 consumers (recorded as patients, and Medicaid enrollees), and 1560 hospitals. Three studies were conducted in the USA and one study was conducted in Canada. Farley 2002a took place in Iowa and Farley 2002b in New Jersey; both studies were set in health plans. Romano 2004 was set in hospitals in California and New York. Tu 2009 was set in hospitals in Canada.

Farley 2002a conducted their study in 35 of the 99 Iowa counties. These counties represented 60% of the total Iowa Medicaid population. The study included MediPass and two types of health maintenance organisations (HMOs) that differed in their performance as assessed with CAHPS surveys scores: one high and one low-rated. The counties were subdivided into three health plan options: type I (MediPass and two HMOs), type II (MediPass and one HMO with a high rating) and type III (MediPass and one HMO with a low rating). The CAHPS survey measures several dimensions of health plan performance including ratings of health plans, primary doctors and reports of experiences with using a health plan. The ratings are for individual items using response scales ranging from 0 to 10. The reports of experiences are composite scores that are averages of responses to sets of individual items with four-category response options.

Farley 2002b was based on the New Jersey Medicaid programme. There was a mandatory HMO enrolment period for Aid for Dependent Children and other welfare-related beneficiaries in 17 of its 21 counties. In February 1998, 91% of these beneficiaries were enrolled in Medicaid HMOs.

Romano 2004 was based on the California Hospital Outcomes Project (CHOP)In California and the Cardiac Surgery Reporting System (CSRS) in New York. Trends in hospital volumes for certain diagnoses after publication of report cards were evaluated. In California the CHOP report published in 1993 evaluated acute myocardial infarction (AMI) mortality at 394 hospitals, complications after lumbar discectomy at 344 hospitals, and complications after cervical discectomy at 277 hospitals. In New York, the CSRS report evaluated 30 hospitals in December 1992 and 31 hospitals in December 1993 and June 1995.

In Canada, Tu 2009 evaluated the public release of performance data of 12 process-of-care indicators for AMI and six indicators for congestive heart failure (CHF) in 86 hospitals. The hospitals were categorised by either early (2004) or delayed (2005) feedback of a publicly released report card about their baseline performance. The Canadian Cardiovascular Outcomes Research Team and the Canadian Cardiovascular Society developed the indicators.

Excluded studies: In total, we excluded 47 studies after assessing full copies of the papers. The main reasons for exclusion were: design (study was not a ClRCT, ClQ-RCT, CBA or ITS (34)), interventions did not contain process measures, health care outcomes, structure measures, consumer or patient experiences, expert- or peer-assessed measures (18), no objective outcome data were recorded or available for one or both arms (seven), and/or the study was about hypothetical choices (six). We excluded four studies because we were unable to obtain the full-text articles (see 'Characteristics of excluded studies').

Risk of bias in included studies

We included three study designs (ClRCT, ClQ-RCT and ITS) which we rated on different risk of bias items, we applied items as appropriate for the relevant study design.

Allocation concealment (selection bias)—One study (Farley 2002a) provided insufficient information about allocation of concealment to allow judment of the degree of the risk of bias. One study (Farley 2002b) described a non-random method of concealing allocation: research investigators enrolling participants could possibly foresee assignment, therefore there is a high risk of bias. In Tu 2009 a statistician randomised participating hospitals stratified by type of hospital, we rated this as a low risk of bias.

Adequate sequence generation (selection bias)—One study (Farley 2002a) provided insufficient information about the sequence generation for judging the degree of the risk of bias. One study (Farley 2002b) described a non-random method of sequence generation (sequence determined by odd or even case record numbers), so it is possible that selection bias occurred. A third study (Tu 2009) used a random method of sequence generation to assign the hospitals to the early feedback group or the delayed feedback group.

Blinding—Blinding of the participants was impossible because they had to see what they received (Farley 2002a; Farley 2002b). Analysis was based on computerised discharged abstracts, for which participants could not be blinded (Romano 2004). It was also impossible to blind hospitals to their randomisation status (Tu 2009).

Incomplete outcome data—Three studies (Farley 2002a; Farley 2002b; Romano 2004) had complete outcome data for the primary outcomes. The results for the entire sample are presented. In Tu 2009, one of 86 hospitals withdrew from the baseline phase after randomisation, and four withdrew from the follow-up phase, all due to resource constraints, although they did not report a reason for the drop-out. We rated this item as having an 'unclear risk of bias': five hospitals dropped out, and this affected both intervention and control groups.

Selective reporting—We have checked two study protocols, the published reports include the expected outcomes Farley 2002a;Farley 2002b. As far as the other two studies were concerned, (Romano 2004; Tu 2009) we were not able to check whether the publications included the expected outcomes.

Other potential sources of bias—Three studies (Farley 2002a; Farley 2002b; Tu 2009) were free of other bias. The ITS study (Romano 2004) had a potential bias since the collection periods were temporally moved about dependent upon when the hospital became an outlier.

Baseline characteristics—In two studies (Farley 2002a; Farley 2002b) the risk of bias regarding the baseline characteristics is unclear, since they did not report demographic variables for the intervention and control groups. One study (Tu 2009) reported the baseline characteristics across their two groups of hospitals.

Reliable outcome measures—Two studies (Farley 2002a; Farley 2002b) achieved appropriate methods for the outcome measurements, and one study (Tu 2009) did not. The primary outcomes measures were developed in a national team of experts, but the measures were not field-tested. The last step in the validation process was not undertaken, thus the reliability of the measures was impossible to determine.

Protection against contamination—The risk of contamination in one study was unclear (Farley 2002a). In another study, the risk of contamination was low because the enrolling participants received the enrolment materials in their homes (Farley 2002b). It is likely that a few respondents would discuss the CAHPS material with others, but the reality is that the risk of contamination cannot be managed in such cases, simply because of the nature of public reporting. The third study did not provide an explicit statement regarding the methods used to prevent against contamination (Tu 2009). There was extensive media coverage following the release of the baseline performance data for the intervention group. The control group also initiated some quality improvement activities after becoming aware of the release of the performance data, which could indicate that the control group had been affected. As in Farley 2002b, this is difficult to prevent because of the nature of public reporting. However, Tu 2009 might have seen that the extended media coverage would affect the hospitals in the control group.

Intervention independent of other changes—In the ITS study (Romano 2004) it is unclear whether the intervention occurred independently of other changes over time or whether that the outcome was influenced by other confounding variables and events during the study period.

Shape of intervention effect pre-specified—The Romano 2004 study adequately pre-specified the shape of the intervention effect.

Knowledge of the allocated interventions adequately prevented during the study—The Romano 2004 study dealt with the knowledge of the allocated interventions suitably.

Intervention unlikely to affect data collection—The Romano 2004 study appropriately managed the risk of affecting the data collection.

Effects of interventions

See: Summary of findings for the main comparison

Characteristics of interventions—In Farley 2002a, conducted between February and May 2000, the control group received standard enrolment materials by personal mailing post, including items such as Medicaid benefits, instructions about the enrolment process, and available information sources. The experimental group received this standard enrolment material plus the Consumer Assessment of Healthcare Providers and System (CAHPS) report. Health plans were categorised on the basis of their CAHPS performance, defined as high and low-performance plans. The CAHPS measures and report template used bar charts rating the overall health plan, overall healthcare and the personal doctor. Additional charts reported respondents' views on five aspects of service: 'getting needed care', 'getting care without long waits', 'how well doctors communicate', 'courtesy, respect and helpfulness', and 'health plan customer service'. A three-point scale (sometimes/never, usually and always) was used. The Iowa Medicaid programme did not offer additional proactive support to the intervention group participants for making health plan choices.

In Farley 2002b, conducted in March and April 1998, the control group beneficiaries received the standard mailing of Medicaid enrolment materials. The experimental group received the standard enrolment material plus the CAHPS report. Following the CAHPS convention for comparative rating, a three-star rating was used with one star for plans with survey results that scored significantly lower than average, two stars for those that were not significantly different from the average for all other Medicaid plans in New Jersey, and three stars for plans that scored significantly better than average. The participants were asked to choose one HMO and sometimes a primary care case-management plan. The state contracted a private firm to manage the enrolment process and assist participants in choosing their plans. They were able to call a free phone number and ask questions about plans. The contractor also sent 'health benefit co-ordinators' into county welfare offices and the community to assist participants in choosing. The Medicaid office automatically assigned participants who did not make to a health plan by the Medicaid office.

In Romano 2004, report cards were published by agencies in California and New York, reporting on patient outcomes for coronary artery bypass grafting (CABG), acute myocardial infarction (AMI) or postdiscectomy complications. The California data began in 1991 with the California Hospital Outcomes Project. The first report, released in California, December 1993, used a two-category rating and classified hospital mortality for AMI and complication rates for cervical and lumbar discectomy as either 'better' and 'not better' than expected. The second report, released in May 1996, classified hospital mortality for AMI into three categories as 'better', 'worse' or 'neither better nor worse' than expected. The analysis for California was based on the California Patient Discharge Data Set.

The New York Cardiac Surgery Reporting System (CSRS) began in 1989 with the creation of a special data system for cardiac surgery. In New York hospital-specific, risk-adjusted

Europe PMC Funders Author Manuscripts

mortality rates using a three-category classification have been released every 12 to 24 months since December 1990. The analysis for New York is based on the Statewide Planning and Research Co-operative System.

In Tu 2009, conducted between April 1999 and April 2005, the early feedback group (42 hospitals) received their baseline performance data of 12 process-of-care indicators for AMI and six indicators for congestive heart failure (CHF) for internal validation checks. The results were publicly released at a press conference and on the internet in January 2004. The early feedback hospitals were encouraged to develop standardised admitting orders and discharge plans, based on the baseline performance. Baseline performance results of the delayed feedback group (N = 41 hospitals) were publicly released on the internet in September 2005 after internal validation. To determine the effect of the public release and feedback, clinical information was collected from chart reviews during the follow-up (1 April 2004 to 31 March 2005 inclusive 15,997 patients) and compared with the baseline performance data (1 April 1999 to 31 March 2001 inclusive 20,039 patients). The primary outcome measures were defined as being the difference in the mean hospital-specific performance between the two study groups on two composite indicators, i.e. one for AMI and one for CHF.

Main outcome measures

Interventions targeting improvement through selection: changes in healthcare utilisation decisions of consumers or *healthcare providers* : In Farley 2002a 22.6% of the participants switched from the default health plan to another health plan. Participants in the type I counties with three plan choices were less likely to switch (19.9%) than those in the type II or type III counties with only two choices (25.4% overall). Availability of CAHPS data had no effect on the switching rate; individuals not receiving information moved from lower to higher quality plans as often as those who did receive CAHPS data (Table 2).

Farley 2002b did not find any significant differences between the plan choices of the enrollees in the intervention and control groups. Sixty-eight percent of the intervention group and 69% of the control group chose a plan. The standardised CAHPS rating for those who chose a plan were -0.03 for the intervention and 0.03 for the control groups; 28% and 27% respectively chose the dominant HMO. For those not selecting the dominant HMO, the standardised CAHPS ratings of the selected plan were 1.80 and 1.73 respectively (Table 3).

Romano 2004 estimated time series models using ordinary least squares (OLS) data from the states of New York and California. They re-analysed the California data with autoregressive integrated moving average (ARIMA) methods. In autoregressive models there were no clear patterns of effect developed between AMI report cards and subsequent hospital volume for either AMI or related AMI conditions. There was a small and temporary increase in volume in low complication rate hospitals for lumbar discectomy. Romano 2004 only report OLS results from New York because autocorrelation was minimal in that state. The study also found a significant increase in CABG volume for low-mortality hospitals in New York within the first month after publication and a significant decrease in volume for high-mortality outliers in the second month after release of the information (Table 4 and Table 5).

Interventions targeting improvement through changes in care: objectives measures of provider performance: Tu 2009 did not find significant differences in either the composite AMI indicator (absolute change 1.5%; 95% CI −2.2% to 5.1%; P = 0.43) (Table 6) or composite CHF indicator (absolute change 0.6%; 95% CI −4.5% to 5.7%; P = 0.81) (Table 7) in the early feedback group compared with the delayed feedback group. Regarding individual process-of-care indicators, one of the 12 for AMI and one of the six for CHF improved significantly in the early feedback group (Table 6 and Table 7). The AMI 30-days mortality rate was significanty lower in the early feedback group than in the delayed feedback group (absolute change −2.5%; 95% CI −0.1% to −4.9%; P = 0.045), while the one-year mortality rates of the early, and delayed feedback groups were comparable. The 30-days and one-year CHF mortality rates did not differ significantly. In addition to the release of a public report card, there was a hospital survey. The early feedback group initiated more quality improvement activities in response to the publicly released report card (for AMI 73.2% versus 46.7%; P = 0.003 and for CHF 61% versus 50%; P = 0.04).

Other outcome measures—Awareness, attitude, views and knowledge of performance data and cost data were not reported in three of the included studies. Farley 2002b reported secondary outcomes as a result of a survey. Farley 2002b used a ratio of 3:1 (report versus no report) to send the survey. They had problems with differential sampling and response rates: therefore the data become difficult to interpret. We decided to exclude these results and so do not report these outcomes.

Data syntheses—We have summarised the outcome data extracted from papers in a narrative format in the section ‘Characteristics of interventions’. We did not synthesise any quantitative data.

DISCUSSION

Summary of main results

In four studies interventions consisting of either direct (mailed) or indirect (internet) release of performance data were focused on changes in selection or changes in care. For changes in selection, from two studies we found no impact on choice of health plan in Medicaid populations. From one study there was a small effect of the public release of mortality and complication data on patient volumes for CABG surgery, and lumbar discectomy; however, these effects did not persist for more than two months after each release. There were no effects of releasing mortality data on patient volumes for AMI outcomes. The changes of care were evaluated for the indirect release of performance data for patients with AMI and CHF in one study. There were no effects on composite process-of-care indicators for either condition, but there were some improvements in individual process-of-care indicators for both AMI and CHF. There was also an effect on 30 day AMI mortality rates, and more quality improvement activities were initiated in response to the publicly-released report cards.

Overall completeness and applicability of evidence

The three studies that took place in the USA involved only a small proportion of the numerous major reporting systems available. We included one new study from Canada (Tu 2009) that was published after the last systematic reviews by Shekelle 2008, Fung 2008 and Faber 2009. We excluded many of the more recent studies because they did not have a rigorous study design or did not report the defined primary outcome measures.

Regarding overall completeness, we conclude that evaluations of public reporting system are scarce. Only a few current reporting systems have been subjected to scientific evaluation to determine the effects of public disclosure of quality information in various stakeholders (Shekelle 2008; Smith 2008). Studies that compare different reporting systems are lacking, as are studies of purchaser behaviour.

Despite evidence that secondary outcome measures (e.g. awareness, attitude, knowledge of performance data) are crucial since public reporting can only change behaviour if the target population (healthcare consumers, providers or purchasers of care) understand the provided information (Hibbard 2010), these measures are lacking in the included studies. Because of that it is difficult to explain the lack of effect. Faber 2009 demonstrated that effect of performance data was higher for those who understand the information. Damman 2011 showed that comparative performance information is complex, and consumers had difficulties in interpreting and using performance data.

One type of performance information included in our studies was about patients experiences (CAHPS), items e.g. regarding doctor-patient communication, long waits, respect. Other included types of performance information were mortality, and complication data. Patient-Reported outcome measures (PROMs) were not included, nor was performance information about services. Mortality and complication data were included, but only for two conditions (AMI, CHF) and two surgical procedures (CABG, discectomy).

Quality of the evidence

The quality of the evidence in this review appears to be low based on the analysis with the Grades of Recommendations Assessment, Development and Evaluation (GRADE) system. We downgraded the quality of the evidence for the outcomes due to some concerns with risk of bias in the studies, loss to follow-up, and very sparse data.

There is one more source of concern regarding the quality of the included studies: we did not have access to the complete study protocols for two of the studies, so we could not judge the risk of selective reporting definitively.

The issue of contamination is difficult to tackle for a public reporting intervention because it is often impossible to prevent control groups from seeing information that is publicly available on websites and in the media. Control of exposure can be gained if the information is only posted to consumers personally or if the control and intervention groups are geographically separated. In one study (Tu 2009), there was extensive media coverage when hospitals in the early feedback group received their baseline performance data. A survey

among hospitals in the delayed feedback group confirmed that these control hospitals were affected by the release of performance data in the early feedback group.

Potential biases in the review process

Although our search was comprehensive, we cannot exclude the possibility of having missed relevant studies. We were unable to retrieve and assess four possibly relevant studies in full text. Two review authors independently examined all the references we found in our search. Two review authors independently extracted detailed data and assessed the risk of bias and a third review author settled any disagreements. We did this to exclude bias in the review process.

Agreements and disagreements with other studies or reviews

There are three relevant publications: an article by Kolstad 2009 and two systematic reviews by Faber 2009 and Fung 2008. Our conclusion agrees with those of Kolstad 2009 and Faber 2009; we do not know the extent to which quality reporting leads to improvement of health care quality. We also agree with the conclusion of Fung 2008; despite the existence of major public reporting systems, we lack rigorous evaluations of the effects of these systems.

AUTHORS' CONCLUSIONS

Implications for practice

The results of this review do not enable us to make any strong recommendations for practice. Whilst performance data may be publicly released for many reasons, we cannot conclude from the limited evidence whether disclosure of performance information can reliably change the behaviour of consumers, providers, purchasers or professionals.

Implications for research

In order to understand the effectiveness of the public release of performance data, we need more longitudinal studies with robust evaluation designs and, in particular, studies that test for delayed or cumulative effects with continuing measurements. To improve our insight into the current and potential impacts of public reporting, we need to evaluate a variety of reporting systems in the USA and other countries.

As the lack of effect might be due to a missing of actual exposure to performance data, a specific implication for future studies targeting the consumer's choice behaviour is that the intervention group (i.e. those provided with performance data) should actually read and understand the performance data. Additional interventions might enhance the impact on consumers with limited health literacy in the intervention group (Hibbard 2007). Studies targeting improvements effected by changes in care might benefit from baseline performance data for the intervention group that is released repeatedly instead of only once.

Berwick's model suggests that public release of performance data may improve quality of care by means of a pathway of change or selection (Berwick 2003). The studies we included focused on either one or the other of these pathways exclusively. We suggest a study design that combines the two pathways to assess the relationship between them.

Europe PMC Funders Author Manuscripts

A basic assumption underlying the provision of report cards is that provider choice is a rational decision. In other words, consumers prefer the healthcare provider or health plan rated as the best. Evidence that confirms this assumption is limited (Faber 2009; Kolstad 2009). However, several factors that influence the choice of consumers are known, such as established relationships with local physicians, health plans (Schwartz 2005; Hibbard 2009), hospitals, distance, and opinions of friends, and family (Harris 2008; The King's Fund 2010). Future studies should address the range, and relative impact of factors such as these.

Acknowledgments

We would like to acknowledge the authors of the original protocol, Phil Alderson and Sandy Oliver, who formulated the idea for this review in 2003. We would also like to acknowledge Craig Ramsay (EPOC Statistical Editor), and Jan Ogaard-Jensen (Norwegian Knowledge Centre for the Health Services) for their statistical assistance. We would like to acknowledge Fiona Beyer (Newcastle University) and Alice Tillema (Radboud University Nijmegen Medical Centre) for their help with the search strategy. Final, we thank the reviewers Donna Farley, Denise O'Connor, and Phil Anderson.

SOURCES OF SUPPORT

Internal sources

- University of Newcastle upon Tyne, UK.

External sources

- UK National Institute for Health Research Cochrane Programme Grant, UK.

SUMMARY OF FINDINGS FOR THE MAIN COMPARISON

Effect of the release of performance data for healthcare consumers or healthcare providers

Patient or population: Patients: treated for AMI or CHF or given CABG or discectomysurgery, and Medicaid enrollees

Settings: Health plans or hospitals

Intervention: Release of performance data (risk adjusted outcomes, consumer reports (CAHPS), composite and individual indicators)

Comparison: No release of performance data

Outcomes	Impact	No of Participants (studies)	Quality of the evidence (GRADE)
Public release of performance data targeted at changes in healthcare utilisation decisions of consumers or healthcare providers			
Changes in selection	In two RCTs there was no significant impact on health plan selection by consumers. In one ITS, across 2 states, the outcomes and complications of 1 medical, and 3 surgical procedures was evaluated. A small and temporary effect (2 months) on patient volume was found for 1 surgical condition in each state	18000 [a] 1474 hospitals (2 RCTs, 1 ITS)	⊕⊕◯◯ **low** [1]
Public release of performance data targeted at improvement through changes in care			

Outcomes	Impact	No of Participants (studies)	Quality of the evidence (GRADE)
Changes in care	There was no effect on two composite measures of heart disease outcome. There was an improvement in two of 18 individual indicators. AMI 30 days mortality rates were improved. The intervention group initiated more quality improvement initiatives	15997 86 hospitals (1 RCT)	⊕⊕◯◯ **low** [2]

GRADE Working Group grades of evidence

High quality: Further research is very unlikely to change our confidence in the estimate of effect.

Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.

Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.

Very low quality: We are very uncertain about the estimate.

[a] Number was based on two studies, the number of participants in the third study was unclear

[1] Moderate risk of bias in study, inconsistent effect across clinical outcomes, sparse data

[2] Some concern with loss to follow-up, very sparse data.

Coronary artery bypass grafting (CABG); acute myocardial infarction (AMI); congestive heart failure (CHF); Consumer Assessment of Healthcare Providers and System (CAHPS) information (CAHPS); Randomised Controlled Trial (RCT); Interrupted time series (ITS) study; Grades of Recommendations Assessment, Development and Evaluation (GRADE)

CHARACTERISTICS OF STUDIES

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

Farley 2002a

Methods	**Design** ClRCT **Unit of allocation** New cases (household units) were randomly assigned to either an experimental or control group. This random assignment was independent of case size, county of residence and initial plan assignment. The Iowa Medicaid office supplied the authors with data files for the full sample of new beneficiaries **Unit of analysis** Medicaid beneficiaries. All analyses were corrected for clustering of beneficiaries within cases using Huber/White corrections **Sample size calculation** Not done. Statistical significance was assessed at the 0.05 level
Participants	**The number randomised into the trial** 13,077 new beneficiaries in 7016 cases; 6515 beneficiaries in the control group, 6562 beneficiaries in the intervention group **Characteristics of participating beneficiaries** 72% of ‘cases’ (family or household units) had 2 or more members with an average of 2.8 beneficiaries per case **Age of beneficiaries** Unclear **Gender of beneficiaries** Unclear **Ethnicity** Unclear **Characteristics of participating health plans:** Two Health Maintenance Organisations (HMOs) under contract with the Medicaid programme and 1 primary care case management plan (MediPass). One was the lower-scoring HMO and the other was the higher-scoring HMO **Setting**

	Iowa, USA. Within the state, counties were divided into 3 groups: those that only contained Medicaid plans and other low-rated plans; those that only contained Medicaid plans and other high-rated plans; and those that contained Medicaid plans and both lower and higher-rated plans
Interventions	**Intervention group** Standard enrolment materials and Consumer Assessment of Healthcare Providers and Systems (CAHPS) report delivered by personal mailing to beneficiaries **Duration of intervention** February 2000 to May 2000 **Intervention deliverer** The Iowa Medicaid office mailed beneficiaries a packet health plan enrolment materials that include items such as a plan enrolment form **Control group** Standard enrolment materials delivered by personal mailing to beneficiaries **Source of funding for study** Co-operative agreement 5U18HS09204-05; the Agency for Healthcare Research and Quality and the Center for Medicare and Medicare Services
Outcomes	**Main outcome** For each of the other 2 counties the authors performed a multinomial logistic regression in which for each beneficiary the outcome took on the values 1 = stay in assigned HMO; 2 = switch to other HMO; 3 = switch to MediPass
Notes	The star charts in the CAHPS report were based on a HMO's performance. The bar charts included 3 charts with ratings of the health plan, healthcare and personal doctor. Five charts were included of service by the providers or health plan

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	The new cases enrolled during the study period were randomly assigned to an experimental or control group. This random assignment was independent of case size, county of residence and initial plan assignment (p. 326)
Allocation concealment (selection bias)	Unclear risk	The new cases enrolled during the study period were randomly assigned to an experimental or control group (p. 326)
Incomplete outcome data (attrition bias) All outcomes	Low risk	"Medicaid office supplied us with data files for the full sample of new beneficiaries (p. 328)" The full sample consisted of 13,077 new beneficiaries in 7016 cases. Results from all are presented in Table 1 (p. 330)
Selective reporting (reporting bias)	Low risk	Only one outcome which was measured in 100% of all cases/new beneficiaries (p. 330)
Other bias	Low risk	Only one outcome which was measured in 100% of all cases/new beneficiaries (p. 330)
Adequate blinding of participants, personnel and outcome assessors?	Unclear risk	"The Iowa Medicaid office supplied us with data files for the full sample of new beneficiaries. The data identified the initially assigned (default) plan, the final plan (different from the default plan only if an active choice was made), whether CAHPS reports were mailed, the county of residence, and the number of beneficiaries in each case (case size)" (p. 328)
Protection against contamination	Unclear risk	The authors claim that "the mail-based administration of the process, including the distribution of CAHPS reports, effectively isolated beneficiaries in the control group (those not receiving CAHPS reports) from exposure to CAHPS information, be it from other beneficiaries outside their household, benefits counsellors, or media information" (p. 327). This statement has not been verified
Baseline characteristics similar?	Unclear risk	Demographic variables could not be measured directly (p. 329) and as such, were not reported for the intervention and control groups."

Reliable outcome measurements	Low risk	The primary outcome measure, i.e. plan choice, was extracted from an automated system (P328). No reliability measure for the procedure was reported

Farley 2002b

Methods	**Design** ClQ-RCT **Unit of allocation** New “cases” (family or household units) were assigned to either an experimental or control group, based on whether the last digit of the case was odd or even. The New Jersey Medicaid office supplied the authors with data files for the full sample of all new 5217 enrollees **Unit of analysis** New Medicaid cases; cases are the family units that qualify for Medicaid coverage. Medicaid-eligible family units that include an adult are referred to as adult cases, and those in which only children are Medicaid-eligible are child cases **Sample size calculation** No justification for the size of the overall sample
Participants	**The number randomised into the trial** 5217 cases. Control: 2, 568 cases. Intervention: 2, 649 cases **Characteristics of participating cases** Cases are the family units that qualify for Medicaid coverage. Medicaid-eligible family units that include an adult are referred to as adult cases, and those in which only children are Medicaid-eligible are child cases. New Jersey requires all members of each Medicaid case to enrol in the same HMO **Age** Unclear **Gender** Men and women **Ethnicity** Whole population **Self reported health status** All health statuses **Characteristics of participating health plans** Medicaid health plans in the state of New Jersey. The Medicaid program has a form of mandatory (auto-assignment) or voluntary managed care programme, which includes one or more HMOs or (sometimes) a primary care case management plan **Setting of care** Medicaid health plans choice between one or more health maintenance organisations (HMOs) and sometimes a primary care case management plan **Country** USA, state of New Jersey
Interventions	**Intervention group** Standard enrolment materials and Consumer Assessment of Health Plans Study (CAHPS) report delivered by personal mailing **Duration of intervention** A 3 week period from 25 March to 15 April 1998 **Intervention deliverer** The New Jersey Medicaid office published a 7-page brochure “choosing an HMO” that compared the Medicaid HMO with respect to the consumer ratings and experiences reported in the CAHPS survey **Control group** Standard enrolment materials delivered by personal mailing **Source of funding** Co-operative agreements No. 5U18HS09204-05; Consumer Assessment of Health Plans Study (CAHPS) from the Agency for Healthcare Research and Quality (AHRQ)
Outcomes	**Main outcome** Auto assignment rates Plan choices
Notes	The star chart in CAHPS report were based on a HMO’s performance compared to the average in every county of residence. The counts ranged from 20 to 29 stars. The resulting standardized CAHPS ratings ranged from −8.40 (well below the average) to 6. 26 (well above the county average)

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	High risk	Quote: "Based on whether the last digit of the case ID was odd or even, half the cases were randomly assigned to an experimental group and half were assigned to a control group" (p. 989)
Allocation concealment (selection bias)	High risk	Allocation concealment was based on case ID number, therefore research investigators enrolling participants could possibly foresee assignment
Incomplete outcome data (attrition bias) All outcomes	Low risk	Quote: "The analysis of the overall effects of CAHPS included the entire April 1998 sample of enrollees, and is therefore not subject to non-response bias" (p.1002)
Selective reporting (reporting bias)	Low risk	Results from all hypotheses listed in methods section are reported
Other bias	Low risk	-
Adequate blinding of participants, personnel and outcome assessors?	Unclear risk	No details were provided about blinding: Quote: "The New Jersey Medicaid office supplied us with a data file that identified plan choices, auto-assignment and demographics for the full sample" (p. 990)
Protection against contamination	Low risk	No information was reported. Quote: "about half the cases mailed a CAHPS report said they received and read the report" (p.996)
Baseline characteristics similar?	Unclear risk	Demographic characteristics were not reported
Reliable outcome measurements	Low risk	Quote: "The primary outcome measure, i.e. plan choice, was extracted from an automated system" (p. 994). No reliability measures for the procedure were reported

Romano 2004

Methods	**Design** Interrupted time series (ITS) **Data analysis** The data were analysed using 2 different approaches: 1) Ordinary least squares (OLS). This method does not take into account the interdependency of subsequent measurements. It is, however, adjusted for several independent variables: state-wise hospital volume for same condition or procedure, hospital, monthly volume before publication of report card for each hospital, interaction between state volume and hospital, unrelated volume in each hospital. The method is used as follows: • based on model predict volume for hospital for each of the 12 months after publication of report card; • model volume based on data preceding publication of report card; • aggregate predicted volumes from all hospitals in same performance category; • estimate 95% confidence intervals (CI) for predictions; • if actual volume falls outside CI, then significant effect of report; 2) Autoregressive (ARIMA) methods with P = 1, q = 0 and the same main effects. As a result of software limitations, the authors were unable to include 2-way interactions. Second-order autoregressive models were also tested but generated very similar results. Because the residual autocorrelations were not significant (P >

	0.10) for all but 1 model, the authors did not estimate moving average models. The autoregressive and OLS results differed, especially for AMI volume, so both sets of numbers were reported to demonstrate the sensitivity of our results to different statistical methods. Autoregressive models adjust for observed correlations, in some analyses, between current and previous volume changes; however only the OLS models adjust for hospital-level interactions
Participants	Patients admitted to hospitals designated as outliers in reports in New York and California **Characteristics of participating patients** The total number of patients with a topic condition or procedure, or related condition or procedure, who were admitted to a specific hospital in a specific calendar month *California - clinical problems:* • Acute myocardial infarction (AMI) • Coronary artery bypass grafting (CABG) (AMI-related) • Percutaneous coronary angioplasty (AMI-related) • Congestive heart failure (AMI-related) • Cervical discectomy (target) • Lumbar discectomy (target) • Back or neck procedures (discectomy-related) • Medical back problems (discectomy-related) • Knee arthroplasty (discectomy-related) • Hip arthroplasty (discectomy-related) *New York - clinical problems:* • AMI • CABG • Percutaneous coronary angioplasty (AMI-related) • Congestive heart failure (AMI-related) **Age** The authors excluded children less than 18 years of age, different groups: < 55 years, 55 to 64 years, 65 to 74 years, > 74 years **Gender** No restriction **Ethnicity:** Black, Hispanic, White **Other characteristics** Patients admitted for psychiatric conditions, injury or poisoning, or rehabilitation were excluded. Also patients transferred from other acute care hospitals were excluded, because transfers generally reflect the capabilities of different facilities, or insurance arrangement, rather than consumer's choice **Characteristics of participating hospitals** Acute care, non-federal hospitals All analyses were limited to hospitals that were included in the report card **Setting of care** Hospital/inpatient **Country** USA, in states of California and New York
Interventions	**Intervention** Annual reports on risk-adjusted outcomes; focused on specific conditions or procedures; the reports incorporate clinical expertise and address regional concerns **Intervention duration** Using hospital months. California: study period 24 months before publication of first report in 1993 and 12 months after. In 1996, 24 months before second report and 7 months after New York: hospital ratings released every 12 to 24 months since December 1990 **Intervention deliverer** Report cards were published by agencies in California and New York **Source of funding** US Agency for Healthcare Research and Quality
Outcomes	**Main outcome**

	Change in the utilisation decisions of consumer, healthcare professional or purchasers
Notes	**Study period** California: First report released in December 1993 classified hospital mortality for AMI and complication rated for cervical and lumbar discectomy as either 'better' or 'not better' than expected. The second report, released in May 1996, classified hospital mortality for AMI as either 'better', 'worse', or 'neither better nor worse' than expected New York: Hospital-specific, risk-adjusted mortality rates and 3-category ratings have been released every 12 to 24 months since December 1990

Risk of bias

Bias	Authors' judgement	Support for judgement
Incomplete outcome data (attrition bias) All outcomes	Low risk	Aggregate data from administrative databases. California analysis was based on the California Patient Discharge Data Set. The New York analysis was based on the Statewide Planning and Research Co-operative System (SPARCS). The number of drop-outs was not applicable here, since the databases and several independent variables were used to predict each hospitals' patient volume after publication of a report card
Selective reporting (reporting bias)	Low risk	All outcomes and results outlined in the Method section are reported in tables and/or text. Results for all primary and secondary objectives are reported
Other bias	High risk	Main analysis based on the assumption of same trend before and after intervention. Difference from predicted values was reported, rather than change in trend and level
Shape of intervention effect pre-specified?	Low risk	Quote: "We predicted what each outlier hospital's volume should have been in each of the 12 months after publication of a report card. These predicted volumes were aggregated for all hospitals assigned to the same risk-adjusted performance category (e.g. higher-than-expected AMI mortality) in that report card". (p.371)
Intervention is independent of other changes?	Unclear risk	It is not sure that the intervention occurred independently of other changes over time or that the outcome was not influence by other confounding variables events during study period
Intervention unlikely to affect / bias data collection?	Low risk	Sources and methods of data collection were the same before and after the intervention
Knowledge of the interventions adequately prevented during the study?	Low risk	Data collection based on administrative database and performed retrospectively
Reliable outcome measurements	Unclear risk	The outcome measures were based on a conceptual framework and hypotheses. Hypotheses were based on a validated assumption (p.368). No clear assessment about the reliability of outcomes measures were reported in the Method section

Tu 2009

Methods	**Design** Cluster-randomised trial **Unit of allocation** Hospitals were randomised to receive either early or delayed feedback of a publicly released report card (p.2331). The randomisation was stratified by type of hospitals (p. 2332). Types of hospitals were classified as teaching hospitals, large community hospitals, small hospitals (p. 2332) **Unit of analysis** Patients treated for acute myocardial infarction (AMI) and congestive heart failure (CHF), taking hospital clustering into account in the analysis **Sample size calculation** The study had 84% power to detect 5% absolute difference on the composite quality indicators. The power calculation assumed a baseline performance rate on each composite indicator of 70% (standard deviation 10%) in each study group, and that there would be a secular improvement of 75% (SD 7.5%) in the composite indicator, independent of the study intervention (p. 2332)
Participants	86 hospital corporations were randomised: 44 hospitals in the early feedback report card release and 42 hospitals in the delayed feedback report card release (Figure 1, p. 2331) **Characteristics of hospitals in the early feedback report card release** At baseline, 5676 patients were admitted with AMI and 5073 patients were admitted with CHF **Characteristics of hospitals in the delayed feedback report card release** At baseline, 5070 patients were admitted with AMI and 4220 patients were admitted with CHF **Setting** The study was conducted in Ontario, Canada. All 130 acute hospitals were assessed for eligibility and 86 hospitals were included
Interventions	**Baseline assessment** At each participating hospital a target sample of 125 charts (or all patients if < 125 patients were treated) for patients receiving care for AMI and/or CHF between 1 April 1999 and 31 March 2001 was abstracted. The baseline performance was based on a set of 12 process of care indicators for AMI and 6 indicators for CHF **Early feedback report card release** The hospitals received their baseline performance data in October 2003 for internal validation and the results were publicly released at a press conference and on the Web in January 2004 (p. 2332) **Duration of intervention** January 2004 to 1 April 1 2004 (inclusive the follow-up period: January 2004 to 31 March 2005) **Delayed feedback report card release** The hospitals received their baseline performance data in September 2005 for internal validation and the results were also publicly released on the internet in September 2005 (p. 2332). No extensive media or associated press was covered **Duration of intervention** Not applicable, as the delayed feedback group received the intervention after follow-up data were collected **Intervention deliverer** The Canadian Cardiovascular Outcomes Research Team, which is a national team of cardiovascular outcomes researchers from across Canada. The team also was involved, together with the Canadian Cardiovascular Society, in the development and endorsement of the set of quality of care indicators, as used in this study **Source of funding for study** The EFFECT study was supported by a Canadian Institutes of Health Research team grant in cardiovascular outcomes research to the Canadian Cardiovascular Outcomes research Team
Outcomes	**Main outcome** There was no significant improvement in the composite AMI or CHF process-of-care indicator. One out of 12 individual process of care AMI indicators improved significant more in the early feedback group compared to the delayed feedback group. One out of 6 of the individual process of CHF indicators improved significantly more in the early feedback group. Regarding mortality rates, as an outcome indicator, 30-day mortality significantly decreased in the early feedback group for AMI, while 3 other mortality-related measures for AMI and CHF did not change The survey showed that the early feedback group reported significantly more often the start of one or more quality improvement initiatives for AMI care and for CHF care
Notes	-

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	The hospitals were randomly assigned to the early feedback group or the delayed feedback group
Allocation concealment (selection bias)	Low risk	Quote: "This random assignment was stratified by type of hospital and performed by a study statistician" (p. 2332)
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	One hospital withdrew from the baseline phase, after randomisation and 4 withdrew from the follow-up phase, all due to resource constraints (p. 2331). No intention-to-treat analysis was performed. Additional exclusions of patients were not reported, but cannot be verified
Selective reporting (reporting bias)	Low risk	Results from all indicators, individual and composite, are reported as well as the hospital outcome indicators
Other bias	Low risk	
Adequate blinding of participants, personnel and outcome assessors?	High risk	Quote: "It was not possible to blind the hospitals to their status" (p. 2332) Quote: "We could not blind the delayed feedback group to the media coverage and associated publicity surrounding the study results" (p. 2336) Quote: "Patient charts were abstracted by an experienced research nurse" (p.2332), but it is unclear whether or not she was blinded for allocation
Protection against contamination	High risk	Quote: "There was extensive media coverage following the release of the baseline performance for the early feedback hospitals" (p. 2332). The authors mention that "one unanticipated observation" was that several hospitals in the delayed feedback group reported that they also initiated some quality improvement activities after becoming aware of the publicly released early feedback report card" (p. 2336)
Baseline characteristics similar?	Low risk	Table 1 (p. 2331) shows the baseline characteristics of the hospitals and patients. Quote: "The hospitals were well balanced across the 2 groups in terms of clinical characteristics of patients" (p. 2333)
Reliable outcome measurements	High risk	Quote: "The primary outcome measures were a set of national process-of-care quality indicators for AMI and CHF care which were developed and endorsed by the Canadian Cardiovascular Outcomes Research Team" (p. 2332). It is likely that the face validity of the indicators is guaranteed, but no field test was performed. The last step in the validation process was not undertaken and no verdict about the content validity is possible. Therefore, we scores this item as high-risk

Cluster-randomised controlled trial (ClRCT); cluster quasi-randomised trial (ClQ-RCT); interrupted time series (ITS); health maintenance organisations (HMOs); coronary artery bypass grafting (CABG); acute myocardial infarction (AMI); congestive heart failure (CHF); standard deviation (SD); confidence interval (CI); Consumer Assessment of Healthcare Providers and System (CAHPS); autoregressive integrated moving average (ARIMA)

Europe PMC Funders Author Manuscripts

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

Study	Reason for exclusion
Alterras 2000	Unable to retrieve
Beaulieu 2002a	Study design
Beaulieu 2002b	Study design
Bundorf 2009	Study design
Dawson 2007	Study design
Dranove 2008	Study design
Ettinger 2008	Study design/outcome measure
Fanjiang 2007	Study design
Fine 1998	Unable to retrieve
Fong 2008	Study design
Fotaki 2008	Study design
Fox 2001	Study design
Goldstein 2001	Study design
Goss 2006	Study design
Hannan 2003	Study design, controlled before-after, 2 intervention sites, only 1 control site
Harris 2002	Study design
Harris-Kojetin 2007	Outcome measure/types of intervention
Hibbard 1996	Outcome measure
Hibbard 2000	Outcome measure
Hibbard 2001	Unable to retrieve
Hibbard 2002a	Study design
Hibbard 2002b	Outcome measures/hypothetical data
Hibbard 2003	Outcome measures
Hibbard 2005a	Study design
Hibbard 2005b	Study design; 2 intervention and 1 control group/raw data was not reported
Hollenbeak 2008	Study design
Jensen 2004	Study design
Jha 2006	Study design
Jian 2009	Study design
Knutson 1998	Study design
Krupat 2004	Types of intervention/outcome measures, design; 2 intervention groups, 1 control group
Lindenauer 2007	Study design
Mannion 2003	Study design
McCormack 2001a	Outcome measures
McCormack 2001b	Outcome measures
Moscucci 2005	Study design, controlled before-after design; no information reported from the 2 included registries. Not enough information was reported regarding the baseline data
Norem 2004	Study design/outcome measures

Europe PMC Funders Author Manuscripts

Study	Reason for exclusion
O'Connor 1991	Unable to retrieve
Peters 2007	Types of intervention/ outcome measures/hypothetical data
Peters 2009	Study design
Schoenbaum 2001	Outcome measures/hypothetical data
Scott 2006	Study design
Spranca 2000	Outcome measures/hypothetical data
Spranca 2007	Study design/outcome measures
Swaminathan 2008	Study design
Tai-Seale 2004	Study design/interventions/outcome measures
Uhrig 2002	Types of intervention/hypothetical data
Uhrig 2006	Outcome measures/hypothetical data
Wedig 2002	Study design, not enough data point for interrupted time series criteria
Werner 2005a	Study design
Werner 2005b	Study design, a single control and single intervention before and after comparison

DATA AND ANALYSES

This review has no analyses.

Appendix 1. MEDLINE search strategy

MEDLINE (OVID)

Syntax guide

/ - index term (MeSH heading)

exp - explode: includes narrower terms to the index term being exploded

.tw. - text word in title or abstract fields

$ - truncation/wild card: adds no or more characters

? - truncation/wild card: adds no or one character

- truncation/wild card: retrieves alternative single character

adjx - adjacency: required words are adjacent to each other, or within × words of each other

.pt. - publication type

Description of search strategy

1. random$.tw.
2. multicenter study.pt.

Europe PMC Funders Author Manuscripts

3. randomized controlled trial.pt.
4. randomized controlled trial.pt.
5. clinical trial.pt.
6. intervention studies/
7. experiment$.tw.
8. (time adj series).tw.
9. (pre test or pretest or (posttest or post test)).tw.
10. random allocation/
11. impact.tw.
12. intervention?.tw.
13. chang$.tw.
14. evaluation studies/
15. evaluat$.tw.
16. effect?.tw.
17. comparative studies/
18. compar$.tw.
19. 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 or 16 or 17 or 18
20. editorial.pt.
21. letter.pt.
22. comment.pt.
23. 20 or 21 or 22
24. animals/
25. humans/
26. 24 not 25
27. 23 or 26
28. 19 not 27
29. (public release of performance data and healthcare providers).mp. [mp=title, original title, abstract, name of substance word, subject heading word]
30. exp Primary Health Care/
31. exp Hospitals/
32. physicians/

Europe PMC Funders Author Manuscripts

33. health professionals.ab,ti.
34. health personnel/
35. health plans.ab,ti.
36. health plan.ab,ti.
37. insurance.ab,ti.
38. *Physician's practice patterns/
39. *Group Practice/
40. *Institutional Practice/
41. *Private Practice/
42. *Family Practice/
43. *Physicians/
44. *Physicians, Family/
45. *Professional Practice/
46. *Nurses/
47. *Nurse Clinicians/
48. *Nurse practitioners/
49. *Pharmacists/
50. *Pharmacies/
51. *Pharmacy/
52. *Hospitals/
53. (physician$ or GP? or doctor? or general pract$ or prescriber? or group pract$ or institutional pract$ or partnership pract$ or family pract$ or general pract$ or office pract$ or private pract$ or primary pract$ or nurse or nurses).tw.
54. (pharmacist? or pharmacies or pharmacy).tw.
55. hospital?.tw.
56. physiotherapist.mp.
57. midwife.mp.
58. health care centre.mp.
59. dietician.mp.
60. health care provider.mp.
61. *Allied Health Personnel/
62. *Dental Clinics/

Europe PMC Funders Author Manuscripts

63. *Dentists/

64. *Outpatient Clinics, Hospital/

65. general pract$.tw.

66. psychologist.mp.

67. psychiatrist.mp.

68. 35 or 36 or 37

69. or/38-67

70. or/30-35

71. 69 or 70

72. 68 or 71

73. quality assurance, health care/

74. *benchmarking/

75. *“process assessment (health care)”/

76. *“outcome assessment (health care)”/

77. exp Quality Indicators, Health Care/

78. performance outcome.ab,ti.

79. (quality adj2 indicator?).tw.

80. (quality adj (criteria or criterion or standard? or norm)).tw.

81. (performance adj (indicator? or measure? or data or rating)).tw.

82. disclosure/

83. Information Services/

84. report card.ab,ti.

85. quality information.ab,ti.

86. public information.ab,ti.

87. consumer information.ab,ti.

88. patient information.ab,ti.

89. 73 or 74 or 75 or 76 or 77 or 78 or 79 or 80 or 81 or 82 or 83 or 84 or 85 or 86 or 87 or 88

90. exp Consumer Satisfaction/

91. patient preferences.ab,ti.

92. public reporting.tw.

93. consumer reports.ab,ti.

94. decision making.ab,ti.
95. choice behaviour.ab,ti.
96. choice behaviour.ab,ti.
97. exp “Patient Acceptance of Health Care”/
98. ‘provider profiling’.ab,ti.
99. 90 or 91 or 92 or 93 or 94 or 95 or 96 or 97 or 98
100. 28 and 72 and 89 and 99

Appendix 2. EMBASE search strategy

EMBASE (OVID)

Syntax guide

/ - index term (EMTREE heading)

exp - explode: includes narrower terms to the index term being exploded

.tw. - text word In title or abstract fields

$ - truncation/wild card: adds no or more characters

? - truncation/wild card: adds no or one character

- truncation/wild card: retrieves alternative single character

adjx - adjacency: required words are adjacent to each other, or within × words of each other

.pt. - publication type

Description of search strategy

1. exp consumer/ or *consumer health information/
2. patient preferences.ab,ti.
3. *patient attitude/
4. *patient participation/
5. *decision making/
6. *patient decision making/
7. 6 or 4 or 1 or 5 or 3 or 2
8. *total quality management/
9. *performance measurement system/
10. public reporting.mp.

Europe PMC Funders Author Manuscripts

11. *decision making/
12. *outcome assessment/
13. *interpersonal communication/
14. *health care quality/ or exp clinical indicator/ or exp “quality of nursing care”/
15. *quality control/
16. report card.ab,ti.
17. public information.mp.
18. consumer information.mp.
19. 13 or 18 or 10 or 14 or 16 or 9 or 15 or 8 or 11 or 17 or 12
20. random$.tw.
21. multicenter study.mp.
22. randomized controlled trial.mp.
23. randomized controlled trial.mp.
24. clinical trial.mp.
25. intervention studies/
26. experiment$.tw.
27. (time adj series).tw.
28. (pre test or pretest or (posttest or post test)).tw.
29. random allocation/
30. impact.tw.
31. intervention?.tw.
32. chang$.tw.
33. evaluation studies/
34. evaluat$.tw.
35. effect?.tw.
36. comparative studies/
37. compar$.tw.
38. 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 or 35 or 36 or 37
39. editorial.pt.
40. letter.pt.
41. comment.pt.

Europe PMC Funders Author Manuscripts

42. 39 or 40 or 41
43. animals/
44. humans/
45. 43 not 44
46. 42 or 45
47. 38 not 46
48. exp Primary Health Care/
49. exp Hospitals/
50. physicians/
51. health professionals.ab,ti.
52. health personnel/
53. health plans.ab,ti.
54. health plan.ab,ti.
55. insurance.ab,ti.
56. (physician$ or GP? or doctor? or general pract$ or prescriber? or group pract$ or institutional pract$ or partnership pract$ or family pract$ or general pract$ or office pract$ or private pract$ or primary pract$ or nurse or nurses).tw.
57. (pharmacist? or pharmacies or pharmacy).tw.
58. hospital?.tw.
59. physiotherapist.mp.
60. midwife.mp.
61. health care centre.mp.
62. dietician.mp.
63. health care provider.mp.
64. general pract$.tw.
65. psychologist.mp.
66. psychiatrist.mp.
67. exp Group practice/
68. exp benchmarking/
69. exp Institutional practice/
70. exp Physician’s Practice Patterns/
71. exp private practice/

72. exp family practice/
73. exp physicians/
74. exp Physicians, family/
75. exp professional practice/
76. exp nurses/
77. exp nurse clinicians/
78. 67 or 63 or 53 or 71 or 70 or 68 or 48 or 77 or 72 or 65 or 55 or 74 or 50 or 75 or 64 or 57 or 61 or 51 or 58 or 69 or 52 or 59 or 60 or 49 or 56 or 73 or 66 or 76 or 62 or 54
79. 7 and 19 and 78 and 47

Appendix 3. PsycINFO search strategy

(OVID)

Syntax guide

/ - index term (APA thesaurus)

exp - explode: includes narrower terms to the index term being exploded

.tw. - text word in title or abstract fields

$ - truncation/wild card: adds no or more characters

? - truncation/wild card: adds no or one character

- truncation/wild card: retrieves alternative single character

adjx - adjacency: required words are adjacent to each other, or within × words of each other

.pt. - publication type

Description of search strategy

1. random$.tw.
2. multicenter study.tw.
3. randomized controlled trial.tw.
4. clinical trial.tw.
5. intervention studies.mp.
6. experiment$.tw.
7. (time adj series).tw.

8. (pre test or pretest or (posttest or post test)).tw.
9. random allocation.mp.
10. impact.tw.
11. intervention?.tw.
12. chang$.tw.
13. evaluation studies.mp.
14. evaluat$.tw.
15. effect?.tw.
16. compar$.tw.
17. comparative studies.mp.
18. 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 or 16 or 17
19. editorial.tw.
20. letter.tw.
21. comment.tw.
22. 21 or 19 or 20
23. animals.mp.
24. humans.mp.
25. 23 not 24
26. 22 or 25
27. 18 not 25
28. Consumer Satisfaction.mp.
29. patient preferences.ab,ti.
30. public reporting.tw.
31. consumer reports.ab,ti.
32. decision making.ab,ti.
33. choice behaviour.ab,ti.
34. 33 or 32 or 28 or 30 or 31 or 29
35. quality assurance, health care.mp.
36. benchmarking.mp.
37. process assessment.mp.
38. outcome assessment.mp.

Europe PMC Funders Author Manuscripts

39. Quality Indicators.mp.
40. performance outcome.ab,ti.
41. (quality adj2 indicator?).tw.
42. (quality adj (criteria or criterion or standard? or norm)).tw.
43. (performance adj (indicator? or measure? or data or rating)).tw.
44. disclosure.mp.
45. Information Services.mp.
46. report card.ab,ti.
47. quality information.ab,ti.
48. public information.ab,ti.
49. consumer information.ab,ti.
50. patient information.ab,ti.
51. 35 or 50 or 39 or 40 or 36 or 41 or 48 or 47 or 38 or 42 or 49 or 46 or 45 or 37 or 43 or 44
52. 27 and 34 and 51

Appendix 4. CINAHL search strategy

EBSCO

Syntax guide

MH - CINAHL subject heading

MM - CINAHL major subject heading

+ - explode: includes narrower terms to the index term being exploded

TI - word in the title field

AB - word in the abstract field

* - truncation/wild card: adds no or more characters

Nx - adjacency: required words are adjacent to each other, or within × words of each other

PT - publication type

Description of search strategy

1. MM “Clinical Trials”
2. TI control* or AB control*

Europe PMC Funders Author Manuscripts

3. TI random* or AB random*
4. MM “Comparative Studies”
5. TI experiment* OR AB experiment
6. TI time N2 series or AB time N2 series
7. TI impact OR AB impact
8. TI intervention* OR AB intervention*
9. Ti evaluat* OR AB evaluat*
10. TI effect? OR AB effect?*
11. “Pretest-Posttest Design+”
12. “quasi-experimental studies+”
13. or/1-12
14. SO cochrane database of systematic reviews
15. 13 not 14
16. MM “Quality of Health Care” or MM “Quality of Nursing Care”
17. MM “Benchmarking”
18. MM “Process Assessment (Health Care)”
19. MM “Outcome Assessment”
20. TI performance outcome OR AB performance outcome
21. TI quality information OR AB quality information
22. TI patient information OR AB patient information
23. TI consumer information OR AB consumer information
24. TI public information OR AB public information
25. TI public reporting OR AB public reporting
26. TI disclosure OR AB disclosure
27. MM “Quality of Health Care”
28. MM “Consumer Satisfaction”
29. TI patient preferences OR AB patient preferences
30. TI consumer reports OR AB consumer reports
31. TI decision making OR AB decision making
32. TI choice behaviour OR AB choice behaviour
33. Ti ‘provider profiling’ or AB ‘provider profiling’
34. TI report card or AB report card

35. S16 or S17 or S18 or S19 or S20 or S21 or S22 or S23 or S24 or S25 or S26 or S34

36. S27 or S28 or S29 or S30 or S31 or S32 or S33

37. S15 and S35 and S36 (exclude Medline records)

Appendix 5. DARE & CENTRAL search strategy

(Cochrane)

(Primary Health Care) OR (Hospitals) OR (Health Personnel) OR (physicians) OR (Nurses) OR (Professional Practice/) OR (Physician ‘s Practice Patterns) OR (Institution Practice) OR (Nurse Clinicians)OR (Pharmacists) OR (Pharmacy) OR (Physiotherapist) OR (Midwife):ti,ab,kw, in Clinical Trials

AND

(Consumer Satisfaction/ OR patient preferences OR public reporting OR consumer reports OR decision making OR choice behaviour OR choice behaviour OR exp “Patient Acceptance of Health Care”/ OR ‘provider profiling’), in Clinical Trials

AND

(quality assurance, health care/) OR (benchmarking) OR (Quality Indicators, Health Care/) OR (disclosure/) OR (Information Services/) OR (report card.ab,ti.) OR (performance outcome.ab,ti.) OR (“Outcome Assessment (Health Care)”/) in Clinical Trials, in Clinical Trials

Appendix 6. Criteria for full-text screening

Nr:

Review author:

	Yes	No	Doubt	Comments
Design: randomised controlled trials, quasi-randomised trials, controlled before-after studies, interrupted time series				Crucial; score: ‘no’ or ‘doubt’ = exclusion
Types of participants: health care providers and professionals, including organisations e.g. hospitals, practice, patients, health care insurance companies, health plans				All participants are important for us, at least 1 type should be central in the study
Types of intervention: participant is exposed to performance information (see*) • Process measures (e.g. waiting times) • Healthcare outcomes (e.g. mortality) • Structure measure (e.g. presence of waiting rooms) • Patient experiences (like CAHPS) • Expert of peer-assessed measures				

	Yes	No	Doubt	Comments
Types of outcome measures: *Primary outcomes:* • Consumer choice of healthcare provider (public and patients) • Healthcare professional choice of healthcare provider • Purchasers choice of healthcare provider • Objective measures of provider performance, including those that were made public and others that were not • Valid measures of staff morale or behaviour ("valid" defined as having the development of the assessment tool reported in a peer reviewed journal) *Secondary outcomes:* • Awareness, attitude, views, knowledge in all target groups • Costs				There should be a description of at least one primary outcome. Only one secondary outcome measure is insufficient
TEMPORARILY INCLUSION				Hypothetical behaviour = exclusion

Appendix 7. Data collection form

Systematic review: the effectiveness of the public release of performance data in changing consumer, healthcare professional or organisational behaviour

(Comments can be made either at the question itself or on a separate sheet (please specify question number)

Name review author:

Date:

Article:

ID article:

Title:

Authors:

Source + year:

Article found in

☐ MEDLINE

☐ EMBASE

Europe PMC Funders Author Manuscripts

☐ CINAHL

☐ PsycINFO

☐ Cochrane

☐ Other (please specify):

☐ Unclear

Study period

Study design

The design of the study is (state which):

☐ Controlled experimental: randomised controlled trial (RCT)
If the author(s) state explicitly (usually by some variant of the term 'random' to describe the allocation procedure used) that the groups compared in the trial were established by random allocation, then the trial is classified as a 'RCT' (randomised controlled trial).

☐ Controlled experimental: controlled clinical trial (CCT) or quasi-randomised studies
If the author(s) do not state explicitly that the trial was randomised, but randomisation cannot be ruled out, the report is classified as a 'CCT' (controlled clinical trial). The classification 'CCT' is where the method of allocation is known but is not considered strictly random, and possibly quasi-randomised trials. Examples of quasi-random methods of assignment include alternation, date of birth and medical record number.

☐ Controlled experimental: controlled before-after study (CBA)
Involvement of intervention and control groups other than by random process, and inclusion of baseline period of assessment of main outcomes. There are two minimum criteria for inclusion of CBAs in EPOC reviews: at least two intervention sites and two control groups are chosen to be similar in respect of the main outcome measures at baseline. Study and control sites are comparable with respect to dominant reimbursement system, level of care, setting of care and/or academic status.

☐ Uncontrolled observational: interrupted time series (ITS)
A change in trend attributable to the intervention. There are two minimum criteria for inclusion of ITS designs in EPOC reviews: clearly defined point in time when the intervention occurred. At least three data points before and three after the intervention.

Classification of study quality

For all study designs

Quality criteria			☐ **N/A**
a) The objective measurement of performance/provider behaviour or patient (health) outcomes			
☐ Done	☐ Not clear	☐ Not done	Page:
b) Relevant and interpretable data presented or obtained			
☐ Done	☐ Not clear	☐ Not done	Page:

For RCT and CCT

Quality criteria	☐ **N/A**
1. Concealment of allocation (protection against selection bias) *(the unit of allocation was by institution, team or professional and any random process is described explicitly, e.g. the use of random number tables, OR the unit of*	

Quality criteria			☐ **N/A**
allocation was by patient or episode of care and there was some form of centralised randomisation scheme, an on-site computer or sealed opaque envelopes were used).			
☐ Done	☐ Not clear	☐ Not done	
2. Follow-up of professionals *(if outcome measures obtained for 80% to 100% of health professionals randomised in the study, do not assume 100% follow-up unless stated explicitly)*			
☐ Done	☐ Not clear	☐ Not done	
3. Follow-up of patients *(if outcome measures obtained for 80% to 100% of subjects randomised or for patients who entered the trial, do not assume 100% follow-up unless stated explicitly)*			
☐ Done	☐ Not clear	☐ Not done	Page:
4. Blinded assessment of primary outcome(s) * *(if the authors state explicitly that the primary outcome variables were assessed blindly OR the outcome variables are objective, e.g. length of hospital stay, drug levels assessed by a standardised test, medical records used)*			
☐ Done	☐ Not clear	☐ Not done	Page:
5. Baseline measurement *(if performance or patient outcomes were measured prior to the intervention, and no substantial differences were present across study groups)*			
☐ Done ☐ N/A	☐ Not clear	☐ Not done	
6. Reliable primary outcome measures *(if two or more raters with at least 90% agreement or kappa $ 0.8 OR the outcome is obtained form some automated system, e.g. length of hospital stay, drug levels assessed by a standardised test)*			
☐ Done	☐ Not clear	☐ Not done	Page:
7. Protection against contamination *(if allocation was by community, institution or practice and it is unlikely that the control group received the intervention)*			
☐ Done	☐ Not clear	☐ Not done	Page:

* If it is a self administered questionnaire: than it is not blinded

Not clear: contact authors

For CBA

Quality criteria			☐ **N/A**
1. Contemporaneous data collection *(if data collection was conducted at the same time as pre and postintervention periods for study and control activities or sites)*			
☐ Done	☐ Not clear	☐ Not done	Page:
2. Baseline measurement *(if performance or patient outcomes were measured prior to the intervention, and no substantial differences were present across study groups)*			
☐ Done	☐ Not clear	☐ Not done	Page:
3. Baseline characteristics are similar for two intervention sites & two control groups in respect of the main outcome measures			
☐ Done	☐ Not clear	☐ Not done	Page:
i. Characteristics for studies using second site *(if characteristics of study and control providers (or patients) are reported and similar)*			
☐ Done	☐ Not clear	☐ Not done ☐ N/A	Page:
ii. Characteristics for studies using untargeted activities as controls *(if study and control activities are comparable with respect to characteristics of targeted behaviour)*			
☐ Done	☐ Not clear	☐ Not done ☐ N/A	Page:

Quality criteria			☐ N/A
iii. Characteristics for studies using patients as control *(if characteristics of study and control providers (or patients) are reported and similar)*			
☐ Done	☐ Not clear	☐ Not done ☐ N/A	Page:
4. Reliable primary outcome measures *(if two or more raters with at least 90% agreement or kappa $ 0.8 OR the outcome is obtained form some automated system, e.g. length of hospital stay, drug levels assessed by a standardised test)*			
☐ Done	☐ Not clear	☐ Not done	Page:
5. Follow-up of professionals *(if outcome measures obtained for 80% to 100% of health professionals randomised in the study, do not assume 100% follow-up unless stated explicitly)*			
☐ Done ☐ Not done	☐ Not clear		Page:
6. Follow-up of patients *(if outcome measures obtained for 80% to 100% of subjects randomised or for patients who entered the trial, do Not assume 100% follow-up unless stated explicitly)*			
☐ Done ☐ Not done	☐ Not clear		Page:
7. Protection against contamination *(if allocation was by community, institution or practice and it is unlikely that the control group received the intervention)*			
☐ Done ☐ Not done	☐ Not clear		Page:
i. Characteristics for studies using second site *(if characteristics of study and control providers (or patients) are reported and similar)*			
☐ Done ☐ Not done	☐ Not clear		Page:
ii. Characteristics for studies using untargeted activities as controls *(if study and control activities are comparable with respect to characteristics of targeted behaviour)*			
☐ Done ☐ Not done	☐ Not clear		Page:
iii. Characteristics for studies using patients as control *(if characteristics of study and control providers (or patients) are reported and similar)*			
☐ Done ☐ Not done	☐ Not clear		Page:

For ITS

Quality criteria			☐ N/A
1. Clearly defined point in time when the intervention occurred.			
☐ Done	☐ Not clear	☐ Not done ☐ N/A	Page:
2. Protection against secular changes:			
i. Intervention is independent of other changes *(if the intervention occurred independent of other changes in time)*			
☐ Done	☐ Not clear	☐ Not done	Page:
ii. Sufficient data points to enable reliable statistical inference *(if at least 3 data points are recorded before and 3 data points recorded after the intervention)*			
☐ Done	☐ Not clear	☐ Not done	Page:
iii. Formal test for trend *(if formal test for change in trend using appropriate method is reported (e.g. Cook & Campbell '79)*			

Quality criteria			☐ N/A
☐ Done	☐ Not clear	☐ Not done	Page:
3. Protection against detection bias:			
iv. Data collection is identical before and after intervention *(if reported that sources and methods of data collection identical before and after intervention)*			
☐ Done	☐ Not clear	☐ Not done	Page:
v. Intervention unlikely to affect/bias data collection *(if reported that the intervention unlikely to affect data collection directly)*			
☐ Done	☐ Not clear	☐ Not done	Page:
vi. Blinded assessment of primary outcome(s) *(if the authors state explicitly that the primary outcome variables were assessed blindly OR the outcome variables are objective, e.g. length of hospital stay, drug levels assessed by standardised test)*			
☐ Done	☐ Not clear	☐ Not done	Page:
4. Completeness of data set *(if data set covers 80% to 100% of total providers and episodes of care in study area)*			
☐ Done	☐ Not clear	☐ Not done	Page:
5. Reliable primary outcome measure(s) *(if two or more raters with at least 90% agreement or kappa $ 0.8 OR the outcome is obtained from some automated system, e.g. length of hospital stay, drug levels assessed by standardised test)*			
☐ Done	☐ Not clear	☐ Not done	Page:

Risk of bias tables of studies with a separate control group (RCTs, CCTs, CBAs)

Risk of bias - Item	Judgement	Description
Sequence generation	Y/N/?	
Allocation concealment	Y/N/?	
Baseline outcome measurements similar	Y/N/?	
Baseline characteristics similar	Y/N/?	
Incomplete outcome data	Y/N/?	
Knowledge of the allocated interventions adequately prevented during the study	Y/N/?	
Protection against contamination	Y/N/?	
Selective outcome reporting	Y/N/?	
Free of other bias	Y/N/?	

Risk of bias tables of interrupted time series studies

Risk of bias - Item	Judgement	Description
Intervention independent of other changes	Y/N/?	
Shape of the intervention effect pre-specified	Y/N/?	
Intervention unlikely to affect data collection	Y/N/?	
Knowledge of the allocated interventions adequately prevented during the study	Y/N/?	
Incomplete outcome data adequately addressed	Y/N/?	

Risk of bias - Item	Judgement	Description
Selective outcome reporting	Y/N/?	
Free of other bias	Y/N/?	

Characteristics of study

Type of participants

☐ Patients/consumers/providers/purchasers (Medicaid enrollees)

		Page
Number		
Patient/consumer/client	☐ patient ☐ client/consumer	
Gender	☐ male ☐ female ☐ mixed	
Age		
Clinical problem		
Other characteristics	1.1.1.1	

☐ Hospitals

		Page
Number		
Size		
Other characteristics	1.1.1.2	

☐ Referring physicians

		Page
Number		
Gender		
Age		
Clinical specialty		
Referring to	☐ primary care ☐ secondary and/or tertiary care	
Other characteristics		

☐ Purchasers of health care: insurance companies

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

	Page
Private or state?	☐ private ☐ state
HMO	☐ yes ☐ no
Collective	☐ yes ☐ no
Company that buys care for employees	☐ yes ☐ no
Other	

Setting of care

- ☐ General practice
- ☐ Outpatient clinic
- ☐ Community care
- ☐ Hospital/inpatient
- ☐ Disabled/inpatient
- ☐ Elderly
- ☐ Any care setting
- ☐ Other (please specify): Medicaid health plans

County

- ☐ North America, including USA and Canada
- ☐ South America
- ☐ Europe
- ☐ Australia or New Zealand
- ☐ Asia
- ☐ Africa
- ☐ Unclear/not specified

Type of control intervention

- ☐ Usual setting
- ☐ Other efforts on quality improvement

Type of intervention: participant is exposed to performance information

All based on actual data

Content of performance information

☐ Process measures (e.g. waiting times)
☐ Patient outcomes (e.g. mortality)
☐ Structure measure (e.g. presence of waiting rooms)
☐ Patient experiences (like CAHPS)
☐ Expert or peer-assessed measures
☐ Other, specify… (‘report cards’)

Description of intervention for both intervention and control groups

Intervention group:	Control group:
Duration of the intervention:	Duration of the intervention:
Intensity:	Intensity:
Duration of follow-up:	Duration of follow-up:
☐ Not clear	☐ Not clear

Information dissemination

	Intervention	Page	Control	Page
Way of data presentation *(numerical, graphical, pictorial summary, star ratings…)*				
How is information made available to the participants *(personal mailing, journal article, active or passive dissemination)*				
Who disseminated the data?				

Outcome measures

Selection:

☐ Changes in the healthcare utilisation decisions of consumer (public and patients)
☐ Changes in the healthcare utilisation decisions of healthcare professional
☐ Changes in the healthcare utilisation decisions of purchasers
Specify:
☐ Choosing the best health plans
☐ Choosing the best healthcare provider (individual physician)
☐ Choosing the best healthcare provider (organisation)
☐ Referring to high quality care provider (individual physician)
☐ Referring to high quality care provider (organisation)
☐ Other.....

Changes in care:

☐ Objective measures of provider performance, including those that were made public and others that were not
☐ Valid measures of staff morale or behaviour ('valid' defined as having the development of the assessment tool reported in a peer reviewed journal)
☐ Other (for example: number of quality improvement efforts):
.....

Attitudes/knowledge/views/understanding/beliefs etc.:

☐ Awareness of information (recall receiving & seen information)
☐ Comprehension of quality of care information (do they understand the information?)
☐ Knowledge about quality of care ('know who is best')
☐ Believes regarding quality of care information (e.g. trust, usefulness, appreciation)
☐ Costs versus quality consideration
☐ Other.....

Data analyses and results

How is outcome data collected?		
Number of observations for primary outcome measure		
Proportion of subjects of study who participate out of the total number in the sampling frame (response rate)	100%	☐ Not clear *(information is not available)*
Number of drop-outs	☐ Not clear	☐ Not applicable
Reason for drop-out mentioned?	☐ Yes ☐ No	☐ Not clear ☐ Not applicable
Confounder or case-mix correction applied?	☐ Yes ☐ No	☐ Not clear ☐ Partially
If Yes: for which variables was corrected? (for example: hospital size)		
Data analysis technique (s):		
State the main results of the main outcome measure(s), for each group (pre- and post values; intervention and control groups), in natural units (mean, SD, n)		
State the corrected intervention effects (mean, 95% confidence interval, P values)		
Describe the main study conclusion		

1. Were the conclusions made by the author(s) supported by the data and/or the analysis reported in the article?	Yes	No	Can't tell/partially
	☐	☐	☐
2. What of the following is applicable to this study? a)Conclusions inconsistent with results b)Conclusions go beyond the data c)No evidence interpreted as no effect d)Implications for research inconsistent with identified shortcoming	Yes	No	Can't tell/partially
	☐	☐	☐
3.Overall (1-2), how would you rate the methods used to analyse the findings relative to the primary question addressed in the study?	☐ Major limitations ☐ Moderate limitations ☐ Minor limitations		

Comments

HISTORY

Protocol first published: Issue 4, 2003

Review first published: Issue 11, 2011

Date	Event	Description
21 August 2008	New citation required and minor changes	Comments on protocol.
4 April 2008	Amended	Converted to new review format.
12 August 2003	New citation required and major changes	Substantive amendment.

DIFFERENCES BETWEEN PROTOCOL AND REVIEW

In the protocol, we listed types of participants as healthcare providers, which included hospitals, practices and individual health professionals. Patients and other healthcare consumers and purchasers of healthcare are also target groups for the aims and scopes of performance measurements. We added these types to the list of participants, so it should be mentioned here, but should not be considered a change of protocol. We mentioned patients, other healthcare consumers and purchasers in the protocol description of outcome measures, but they were missing in the types of participants. We solved this inconsistency in the review by adding these types of participants.

Since the publication of the protocol, the Effective Practice and Organisation of Care (EPOC) Group has adjusted the definitions for the quality criteria. In the review, we used the latest version of the ‘Risk of bias’ tables to assess the quality of the studies.

References to studies included in this review

Farley 2002a {published data only} . *Farley DO, Elliott MN, Short PF, Damiano P, Kanouse DE, Hays RD. Effect of CAPHS Performance Information on health plan choices by Iowa Medicaid. Medical Care Research and Review. 2002; 59(3):319–36. [PubMed: 12205831]

Farley 2002b {published data only} . *Farley DO, Short PF, Elliott MN, Kanouse DE, Brown JA, Hays RD. Effect of CAPHS health plan performance information on plan choices by New Jersey. Health Services Research. 2002; 37(4):985–1007. [PubMed: 12236394]

Romano 2004 {published data only} . *Romano PS, Hong Zhou MPH. Do well-publicized risk-adjusted outcomes reports affect hospital volume? Medical Care. 2004; 42(4):367–77. [PubMed: 15076814]

Tu 2009 {published data only} . Tu JV, Donovan LR, Douglas SL, 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.

References to studies excluded from this review

Alterras 2000 {published data only} . Alterras TT. Health plan report cards may influence insurers more than consumers: their effect on insurer behavior in Minnesota. Health Care Financing and Organization. 2000; 3(3):1–2.

Beaulieu 2002a {published data only} . Beaulieu ND, Epstein AM. National Committee on Quality Assurance health-plan accreditation: predictors, correlates of performance, and market impact. Medical Care. 2002; 40(4):325–87. [PubMed: 12021688]

Beaulieu 2002b {published data only} . Beaulieu ND. Quality information and consumer health plan choices. Journal of Health Economics. 2002; 21(1):43–63. [PubMed: 11845925]

Bundorf 2009 {published data only} . Bundorf MK, Chun N, Goda GS, Kessler DP. Do markets respond to quality information? The case of fertility clinics. Journal of Health Economics. 2009; 28(3):718–27. [PubMed: 19328568]

Dawson 2007 {published data only} . Dawson D, Gravelle H, Jacobs R, Martin S, Smith PC. The effects of expanding patient choice of provider on waiting times: evidence from a policy experiment. Health Economics. 2007; 16(2):113–28. [PubMed: 16888753]

Dranove 2008 {published data only} . Dranove D, Sfekas A. Start spreading the news: a structural estimate of the effects of New York hospital report cards. Journal of Health Economics. 2008; 27(5):1201–7. [PubMed: 18420293]

Ettinger 2008 {published data only} . Ettinger WH, Hylka SM, Phillips RA, Harrison LH Jr, Cyr JA, Sussman AJ. When things go wrong: the impact of being a statistical outlier in publicly reported coronary artery bypass graft surgery mortality data. American Journal of Medical Quality. 2008; 23(2):90–5. [PubMed: 18245577]

Fanjiang 2007 {published data only} . Fanjiang G, von Glahn T, Chang H, Rogers WH, Safran DG. Providing patients web-based data to inform physician choice: if you build it, will they come? Journal of General Internal Medicine. 2007; 22(10):1463–6. [PubMed: 17653810]

Fine 1998 {published data only} . Fine A. The effect of grading and publicizing the physician groups' performance. Executive Solutions for Healthcare Management. 1998; 1(9):2–3.

Fong 2008 {published data only} . Fong J, Marsh GM, Stokan LA, Sang W, Vinson C, Ruhl L. Hospital quality performance report: an application of composite scoring. Journal of Medical Quality. 2008; 23(4):287–94.

Fotaki 2008 {published data only} . Fotaki M, Roland M, Boyd A, McDonald R, Scheaff R, Smith L. What benefits will choice bring to patients? Literature review and assessment of implications. Journal of Health Service Research Policy. 2008; 13(3):178–84.

Fox 2001 {published data only} . Fox MH, Moore J, Zimmerman M, Hill S, Foster CH. The effectiveness of CAHPS among women enrolling in Medicaid managed care. Journal of Ambulatory Care Management. 2001; 24(4):76–91. [PubMed: 11680241]

Goldstein 2001 {published data only} . Goldstein E, Fyock J. Reporting of CAHPS quality information to medicare beneficiaries. Health Services Research. 2001; 36(3):477–88. [PubMed: 11482585]

Goss 2006 {published data only} . Goss JR, Maynard C, Aldea GS, Marcus-Smith M, Whitten RW, Johnston G, Phillips RC, Reisman M, Kelley A, Anderson RP, Clinical Outcomes Assessment Program. Effects of a statewide physician-led quality-improvement program on the quality of cardiac care. American Hearth Journal. 2006; 151(5):1033–42.

Hannan 2003 {published data only} . Hannan EL, Sarrazin MS, Doran DR, Rosenthal GE. Provider profiling and quality improvement efforts in coronary artery bypass graft surgery: the effect on short-term mortality among Medicare beneficiaries. Medical Care. 2003; 41(10):1164–72. [PubMed: 14515112]

Harris 2002 {published data only} . Harris KM. Can high quality overcome consumer resistance to restricted provider access? Evidence from a health plan choice experiment. Health Services Research. 2002; 37(3):551–71. [PubMed: 12132595]

Harris-Kojetin 2007 {published data only} . Harris-Kojetin LD, Uhrig JD, Williams P, Bann C, Frentzel EM, McCormack L, et al. The "choose with care system" - development of education

materials to support informed Medicare health plan choices. Journal of Health Communication. 2001; 12(2):133–56. [PubMed: 17365356]

Hibbard 1996 {published data only} . Hibbard JH, Sofaer S, Jewett JJ. Condition-specific performance information: assessing salience, comprehension, and approaches for communicating quality. Healthcare Financing Review. 1996; 18(1):95–109.

Hibbard 2000 {published data only} . Hibbard JH, Harris-Kojetin L, Mullin P, Lubalin J, Garfinkel S. Increasing the impact of health plan report cards by addressing consumers' concerns. Health Affairs. 2000; 19(5):138–43. [PubMed: 10992661]

Hibbard 2001 {published data only} . Hibbard JH, Peters E, Slovic P, Finucane ML, Tusler M. Making health care quality reports easier to use. Joint Commission Journal on Quality Improvement. 2001; 27(11):591–604. [PubMed: 11708039]

Hibbard 2002a {published data only} . Hibbard JH, Berkman N, McCormack LA, Jael E. The impact of a CAHPS report on employee knowledge, beliefs, and decisions. Medical Care Research and Review. 2002; 59(1):104–16. [PubMed: 11877876]

Hibbard 2002b {published data only} . Hibbard JH, Slovic P, Peters E, Finucane ML. Strategies for reporting health plan performance information to consumers: evidence from controlled studies. Health Services Research. 2002; 37(2):291–313. [PubMed: 12035995]

Hibbard 2003 {published data only} . Hibbard JH, Stockard J, Tusler M. Does publicizing hospital performance stimulate quality improvement efforts? Health Affairs. 2003; 22(2):84–94. [PubMed: 12674410]

Hibbard 2005a {published data only} . Hibbard JH, Stockard J, Tusler M. It isn't just about choice: the potential of a public performance report to affect the public image of hospitals. Medical Care Research and Review. 2005; 62(2):358–71. [PubMed: 15894709]

Hibbard 2005b {published data only} . Hibbard JH, Stockhard J, Tusler M. Hospital performance reports: impact on quality, market share, and reputation. Health Affairs. 2005; 24(4):1150–60. [PubMed: 16012155]

Hollenbeak 2008 {published data only} . Hollenbeak CS, Gorton CP, Tabak YP, Jones JL, Milstein A, Johannes RS. Reduction in mortality associated with intensive public reporting of hospital outcomes. American Journal of Medical Quality. 2008; 23(4):279–86. [PubMed: 18658101]

Jensen 2004 {published data only} . Jensen J. The ratings game: how important are hospital rankings to consumers? Marketing Health Services. 2004; 24(1):40–5. [PubMed: 15008008]

Jha 2006 {published data only} . Jha AK, Epstein AM. The predictive accuracy of the New York State coronary artery bypass surgery report-card system. Health Affairs. 2006; 25(3):844–55. [PubMed: 16684751]

Jian 2009 {published data only} . Jian W, Huang Y, Hu M, Zhang X. Performance evaluation of inpatient service in Beijing: a horizontal comparison with risk adjustment based on Diagnosis Related Groups. BMC Health Services Research. 2009; 9:72–1. [PubMed: 19402913]

Knutson 1998 {published data only} . Knutson DJ, Kind EA, Fowles JB, Adlis S. Impact of report cards on employees: a natural experiment. Health Care Financing Review. 1998; 20(1):5–27. [PubMed: 10387425]

Krupat 2004 {published data only} . Krupat E, Hsu J, Irish J, Schmittdiel JA, Selby J. Matching patients and practitioners based on beliefs about care: results of a randomized controlled trial. American Journal of Managed Care. 2004; 10(11 Pt 1):814–22. [PubMed: 15623271]

Lindenauer 2007 {published data only} . Lindenauer PK, Remus D, Roman S, Rothberg MB, Benjamin EM, Ma A, et al. Public reporting and pay for performance in hospital quality improvement. New England Journal of Medicine. 2007; 356(5):486–96. [PubMed: 17259444]

Mannion 2003 {published data only} . Mannion R, Goddard M. Public disclosure of comparative clinical performance data: lessons from the Scottish experience. Journal of Evaluation in Clinical Practice. 2003; 9(2):277–86. [PubMed: 12787191]

McCormack 2001a {published data only} . McCormack LA, Garfinkel SA, Hibbard JH, Norton EC, Bayen UJ. Health plan decision making with new Medicare information materials. Health Services Research. 2001; 36(3):531–54. [PubMed: 11482588]

McCormack 2001b {published data only} . McCormack LA, Anderson WL, Uhrig JD, Garfinkel SA, Sofaer S, Terrell SA. Health plan decision making in the Medicare population: results from a national randomized experiment. Health Services Research. 2001; 36(6 Pt 2):133–49. [PubMed: 16148965]

Moscucci 2005 {published data only} . Moscucci M, Eagle KA, Share D, Smith D, De Franco AC, O'Donnell M, et al. Public reporting and case selection for percutaneous coronary interventions: an analysis from two large multicenter percutaneous coronary intervention databases. Journal of the American College of Cardiology. 2005; 45(11):1759–65. [PubMed: 15936602]

Norem 2004 {published data only} . Norem J, Moen MA. The websites of Norwegian hospitals: do they meet national guidelines and patient's expectations? Journal of Telemedicine and Telecare. 2004; 10(5):272–6. [PubMed: 15532132]

O'Connor 1991 {published data only} . O'Connor SJ, Shewchuk RM, Bowers MR. A model of service quality perceptions and health care consumer behavior. Journal of Hospital Marketing. 1991; 6(1):69–92. [PubMed: 10116639]

Peters 2007 {published data only} . Peters E, Dieckmann N, Dixon A, Hibbard JH, Mertz CK. Less is more in presenting quality information to consumers. Medical Care Research and Review. 2007; 64(2):169–90. [PubMed: 17406019]

Peters 2009 {published data only} . Petersen LA, Woodard LD, Henderson LM, Urech TH, Pietz K. Will hypertension performance measures used for pay-for-performance programs penalize those who care for medically complex patients? Circulation. 2009; 119(23):2978–85. [PubMed: 19487595]

Schoenbaum 2001 {published data only} . Schoenbaum M, Spranca M, Elliott M, Bhattacharya J, Short PF. Health plan choice and information about out-of-pocket costs: an experimental analysis. Inquiry. 2001; 38(1):35–48. [PubMed: 11381720]

Scott 2006 {published data only} . Scott IA, Ward M. Public reporting of hospital outcomes based on administrative data: risks and opportunities. Medical Journal of Australia. 2006; 184(11):571–5. [PubMed: 16768665]

Spranca 2000 {published data only} . Spranca M, Kanouse DE, Elliott M, Short PF, Farley DO, Hays RD. Do consumer reports of health plan quality affect health plan selection? Health Services Research. 2000; 35(5 Pt 1):933–47. [PubMed: 11130805]

Spranca 2007 {published data only} . Spranca MD, Elliott MN, Shaw R, Kanouse DE. Disenrollment information and Medicare plan choice: is more information better? Health Care Financing Review. 2007; 28(3):47–59. [PubMed: 17645155]

Swaminathan 2008 {published data only} . Swaminathan S, Chernew M, Scanlon DP. Persistence of HMO performance measures. Health Services Research. 2008; 43(6):2033–49. [PubMed: 18783460]

Tai-Seale 2004 {published data only} . Tai-Seale M. Does consumer satisfaction information matter? Evidence on member retention in FEHBP plans. Medical Care Research and Review. 2004; 61(2):171–86. [PubMed: 15155050]

Uhrig 2002 {published data only} . Uhrig JD, Short PF, Farley P. Testing the effect of quality reports on the health plan choices of Medicare beneficiaries. Inquiry. 2002; 39(4):355–71. [PubMed: 12638711]

Uhrig 2006 {published data only} . Uhrig JD, Harris-Kojetin L, Bann C, Kuo TM. Do content and format affect older consumers' use of comparative information in a Medicare health plan choice? Results from a controlled experiment. Medical Care Research and Review. 2006; 63(6):701–18. [PubMed: 17099122]

Wedig 2002 {published data only} . Wedig GJ, Tai-Seale M. The effect of report cards on consumer choice in the health insurance market. Journal of Health Economics. 2002; 21(6): 1031–48. [PubMed: 12475124]

Werner 2005a {published data only} . Werner RM, Asch DA. The unintended consequences of publicly reporting quality information. JAMA. 2005; 293(10):1239–44. [PubMed: 15755946]

Werner 2005b {published data only} . Werner RM, Asch DA, Polsky D. Racial profiling: the unintended consequences of coronary artery bypass graft report cards. Circulation. 2005; 111(10):1257–63. [PubMed: 15769766]

Additional references

Bardach 2009 . Bardach NS, Cabana MD. The unintended consequences of quality improvement. Current Opinion in Pediatrics. 2009; 21(6):777–82. [PubMed: 19773653]

Bentley 1998 . Bentley JM, Nash DB. How Pennsylvania hospitals have responded to publicly released reports on coronary artery bypass graft surgery? The Joint Commission Journal on Quality Improvement. 1998; 24(1):40–9. [PubMed: 9494873]

Berwick 1990 . Berwick DM, Wald DL. Hospital leaders' opinions of the HCFA mortality data. JAMA. 1990; 263(2):247–9. [PubMed: 2403602]

Berwick 2003 . Berwick D, Jamer B, Coye M. Connections between quality measurements and improvement. Medical Care. 2003; 41(1):130–8.

Brook 1994 . Brook RH. Health care reform is on the way: do we want to compete on quality? Annals of Internal Medicine. 1994; 120(1):84–6. [PubMed: 8250463]

Damman 2010 . Damman OC, Van den Hengel YKA, Van Loon AJM, Rademakers J. An international comparison of Web-based reporting about healthcare quality: content analysis. Journal of Medical Internet Research. 2010; 13(12):e8. [PubMed: 20439252]

Damman 2011 . Damman OC, Hendrik M, Rademakers J, Spreeuwenberg P, Delnoij DMJ, Groenewegen PP. Consumers interpretation and use of comparativeinformation on the quality of health care: the effectof presentation approaches. Health Expectations. 2011:15. **doi: 10.1111/j.1369-7625.2011.00671.x**(first published online). [PubMed: 21838831]

EPOC 2009 . Effective Practice and Organisation of Care Group. EPOC risk of bias guideline. http://epoc.cochrane.org/epoc-resources-review-authorshttp://epoc.cochrane.org/epoc-resources-review-authors

Faber 2009 . Faber M, Bosch M, Wollersheim H, Leatherman S, Grol R. Public reporting in health care: how do consumers use quality of care information? A systematic review. Medical Care. 2009; 47(1):1–8. [PubMed: 19106724]

Fung 2008 . Fung C, Yee-Wei L, Soeren M, Damberg C, Shekelle P. Systematic review: the evidence that publishing patient care performance data improves quality of care. Annals of Internal Medicine. 2008; 148(2):111–23. [PubMed: 18195336]

Giuffrida 1999 . Giuffrida A, Gravelle H, Roland M. Measuring quality of care with routine data: avoiding confusion between performance indicators and health outcomes. BMJ. 1999; 319(7202):94–8. [PubMed: 10398635]

Harris 2008 . Harris, KM.; Beeuwkes Buntin, M.; The RAND Cooperation. Research Synthesis Report. Choosing a healthcare provider: the role of quality information. Robert Wood Johnson Foundation; Princeton: 2008.

Hendriks 2009 . Hendriks M, Spreeuwenberg P, Rademakers J, Delnoij DM. Dutch healthcare reform: did it result in performance improvement of health plans? A comparison of consumer experiences over time. BMC Health Services Research. 2009; 9:167. [DOI: 10.1186/1472-6963-9-167]. [PubMed: 19761580]

Hibbard 1997 . Hibbard JH, Jewett JJ, Legnini MW, Tusler M. Choosing a health plan: do large employers use the data? Health Affairs. 1997; 16(6):172–80. [PubMed: 9444825]

Hibbard 2007 . Hibbard JH, Peters E, Dixon A, Tusler M. Consumers competencies and the use of comparative quality information: it isn't just about literacy. Medical Care Research and Review. 2007; 64(4):379–94. [PubMed: 17684108]

Hibbard 2009 . Hibbard JH. Using systematic measurement to target consumer activation strategies. Medical Care Research and Review. 2009; 66(1):9S–27S. [PubMed: 19052169]

Hibbard 2010 . Hibbard JH, Greene J, Daniel D. What is quality anyway? Performance reports that clearly communicate to consumers the meaning of quality of care. Medical Care Research and Review. 2010; 67(3):275–93. [PubMed: 20093399]

Higgins 2011 . Higgins, JPT.; Green, S., editors. Cochrane Handbook for Systematic Reviews of Interventions. Version 5.1.0. The Cochrane Collaboration; 2011. www.cochrane-handbook.org [updated March 2011]

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].

Kerr 2007 . Kerr EA, Hofer TP, Hayward RA, Adams JL, Hogan MM, McGlynn EA, et al. Quality by any other name? A comparison of three profiling systems for assessing health care quality. Health Services Research. 2007; 42(5):2070–87. [PubMed: 17850534]

Kolstad 2009 . Kolstad JT, Chernew ME. Quality and consumer decision making in the market for health insurance and health care services. Medical Care Research and Review. 2009; 66(Suppl 1):28S–52S. [PubMed: 19029288]

Loeb 2004 . Loeb JM. The current state of performance measurement in health care. International Journal for Quality in Health Care. 2004; 16(Suppl 1):i5–9. [PubMed: 15059982]

Longo 1997 . Longo DR, Land G, Schramm W, Fraas J, Hoskins B, Howell V. Consumer reports in health care. Do they make a difference in patient care?. JAMA. 1997; 278(19):1579–84.

Magee 2003 . Magee H, Davis LJ, Coulter A. Public views on healthcare performance indicators and patient choice. Journal of the Royal Society of Medicine. 2003; 96(7):338–42. [PubMed: 12835446]

Marshall 2000 . Marshall MN, Shekelle PG, Leatherman S, Brook RH. The public release of performance data. What do we expect to gain? A review of the evidence. JAMA. 2000; 283(14): 1866–74. [PubMed: 10770149]

Marshall 2004 . Marshall MN, Romano PS, Davies HT. How do we maximize the impact of the public reporting of quality of care? International Journal for Quality in Health Care. 2004; 16(Suppl 1):i57–63. [PubMed: 15059988]

Moher 1999 . Moher D, Cook DJ, Eastwood S, Olkin I, Drummon R, Stroup DF, QUOROM Group. Improving the quality of reports of meta-analyses of randomised controlled trials: the QUOROM statement. Quality of Reporting of Meta-analyses. Lancet. 1999; 354(9193):1896–900. [PubMed: 10584742]

Mukamel 1998 . Mukamel DB, Mushlin AI. Quality of care information makes a difference: an analysis of market share and price changes after publication of the New York State Cardiac Surgery Mortality Reports. Medical Care. 1998; 36(7):945–54. [PubMed: 9674613]

O'Meara 2005 . O'Meara J, Kitchener M, Collier E, Lyons M, De Billwiller-Kiss A, Simon LP, et al. Case study: development of and stakeholder responses to a nursing home consumer information system. American Journal of Medical Quality. 2005; 20(1):40–50. [PubMed: 15782754]

Parry 1998 . Parry GJ, Gould CR, McCabe CJ, Tarnow-Mordi WO. Annual league tables of mortality in neonatal intensive care units: longitudinal study. International Neonatal Network and the Scottish Neonatal Consultants and Nurses Collaborative Study Group. BMJ. 1998; 316(7149):1931–5. [PubMed: 9641927]

Rixom 2002 . Rixom A. Performance league tables. BMJ. 2002; 325(7357):177–8. [PubMed: 12142293]

Rosenthal 1998 . Rosenthal GE, Hammar PJ, Way LE, Shipley SA, Doner D, Wojtala B, et al. Using hospital performance data in quality improvement: the Cleveland Health Quality Choice experience. Joint Commission Journal of Quality Improvement. 1998; 24(7):347–60.

Schut 2005 . Schut FT, van de Ven WP. Rationing and competition in the Dutch health-care system. Health Economics. 2005; 14(Suppl 1):S59–74. [PubMed: 16161190]

Schwartz 2005 . Schwartz LM, Woloshin S, Birkmeyer JD. How do elderly patients decide where to go for major surgery? Telephone interview survey. BMJ. 2005; 331(7520):821. [DOI: 10.1136/bmj.38614.449016.DE]. [PubMed: 16192286]

Shekelle 2008 . Shekelle, PG.; Lim, Y-W.; Mattke, S.; Damberg, C.; Southern California Evidence-based Practice Centre, RAND Corporation. Does public release of performance results improve quality of care? A systematic review. The Health Foundation; London: 2008. Does public release of performance results improve quality of care? A systematic review.

Sirio 1996 . Sirio CA, McGee JL. Public reporting of clinical outcomes--the data needs of health care stakeholders. American Journal of Medical Quality. 1996; 11(1):S78–81. [PubMed: 8763242]

Smith 2008 . Smith, PC.; Mossialos, E.; Papanicolas, I. Performance measurement for health system improvement: experiences, challenges and prospects. World Health Organization; 2008.

Smith 2009 . Smith, PC.; Mossialos, E.; Papanicolas, I.; Leatherman, S. Performance measurement for health system improvement. Experiences, Challenges and Prospects. Cambridge University Press; Cambridge: 2009. p. 3-726.

The King's Fund 2010 . Dixon, A.; Roberson, R.; Appleby, J.; Burge, P.; Devlin, N.; Magee, H. Patient Choice. How patients choose and how providers respond. The King's Fund; London: 2010.

Werner 2009 . Werner RM, Konetzka RM, Stuart EA, Norton EC, Polsky D, Park J. Impact of public reporting on quality of postacute care. Health Services Research. 2009; 44(4):1169–87. [PubMed: 19490160]

* *Indicates the major publication for the study*

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

PLAIN LANGUAGE SUMMARY

Public release of performance data in changing the behaviour of healthcare consumers, professionals or organisations

The objective of the review was to determine whether publicly releasing performance data changes the behaviour of healthcare consumers, professionals, providers and purchasers in a way that improves performance and quality of care. Although it is increasingly common to release information about performance in health care, its effectiveness in changing behaviour has not been determined. We searched the scientific literature for studies comparing the public release of performance data using different public release approaches. These approaches varied in type, content and presentation formats. The aim of the studies was to assess how and to what extent data release led to changes in care. We examined the effects of interventions by targeting the behaviour of healthcare consumers, professionals, providers and purchasers. We also examined the improvement effected by changes in care.

We found four relevant studies containing more than 35,000 consumers, and 1560 hospitals. Two were in health plan settings and the other two were in hospital settings. One cluster-randomised controlled trial and one cluster quasi-randomised trial reported no evidence that information from the Consumer Assessment of Healthcare Providers and System influenced the health plan choice in a Medicaid population. The interrupted time series study found a small positive effect of publishing data on patient volumes for coronary artery bypass grafting and low-complication outliers for lumbar discectomy, but these effects did not persist longer than two months after each public release. No effects on patient volumes for acute myocardial infarction were found. One cluster-randomised controlled trial studied performance and quality of care after the public release of performance data for patients with acute myocardial infarction and congestive heart failure. No effects were found for the composite process-of-care indicators for either condition, but there were some improvements in the individual indicators for acute myocardial infarction and congestive heart failure and in the acute myocardial infarction mortality rates. Further, more quality improvement activities were initiated as a result. No secondary outcomes (awareness, attitude, views and knowledge of performance data and costs) were reported. On the basis of this rather sparse evidence of low quality, we cannot draw any firm conclusions about the effects of the public release of performance data on behaviour or on improvements in the performance and quality of health care.

Europe PMC Funders Author Manuscripts

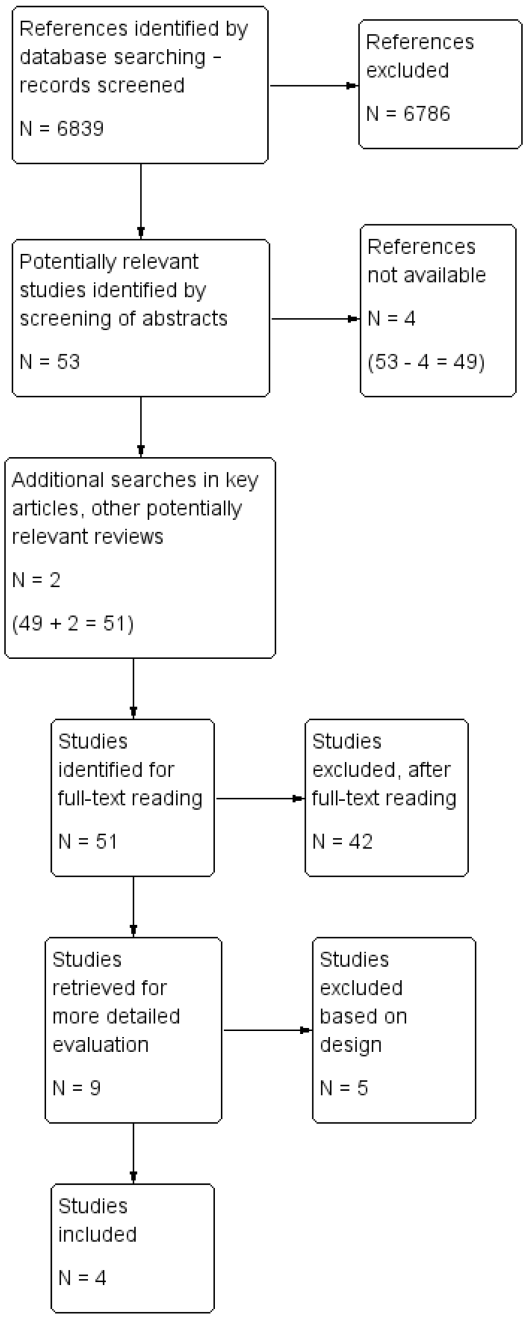

Figure 1.
Flowchart for Study Selection

Europe PMC Funders Author Manuscripts

Table 1

Characteristics of settings and consumers (recorded as patients and Medicaid enrolees)

Study	Methods	Patients/consumers	Setting of care
Farley 2002a	**Design:** ClRCT **Unit of allocation:** new cases (household units) **Power calculation:** not done	**Medicaid beneficiaries:** 13077 **Age:** unclear **Gender:** unclear	**Health plans:** HMOs
Farley 2002b	**Design:** ClQ-RCT **Unit of allocation:** new cases (household units) **Power calculation:** unclear	**Medicaid beneficiaries:** 5878 **Age:** unclear **Gender:** men and women	**Health plans:** HMOs
Romano 2004	**Design:** ITS **Unit of allocation:** not applicable **Power calculation:** unclear	**Patients:** treated for CABG in New York, and for AMI and post discectomy complications in California Number of patients: unclear **Age:** children younger than 18 years were excluded **Gender:** men and women	**Hospitals:** non-federal hospitals
Tu 2009	**Design:** ClRCT **Unit of allocation:** hospitals **Power calculation:** the study had 84% power to detect 5% absolute difference in the composite quality indicators. The assumptions were a baseline performance rate of 70% (SD 10%) for each composite indicator in each study group, and a secular improvement of 75% (SD 7.5%) in the composite indicator, independent of the study intervention	**Patients:** 15997 patients treated for AMI Or CHF **Age:** no restriction **Gender:** men and women	**Hospitals:** teaching, community or small

Cluster-randomised controlled trial (ClRCT); cluster quasi-randomised trial (ClQ-RCT); interrupted time series (ITS); health maintenance organisations (HMOs); coronary artery bypass grafting (CABG); acute myocardial infarction (AMI); congestive heart failure (CHF); standard deviation (SD).

Table 2

Estimated effects of CAHPS information on enrolment choices by new beneficiaries enrolled in the Iowa Medicaid programme

	Percentages (Unadjusted Frequencies)				
Type I counties	**Control (No CAHPS[1])**	**Intervention (CAHPS)**	**Odds ratio for the CAHPS Group**	**Lower CI[2]**	**Upper CI**
Assigned to high-rated HMO	N = 1717	N = 1693			
Stayed in HMO[3]	84.0%	85.7%			
Switched to MediPass[3]	13.2%	10.6%	0.80	0.58	1.09
Switched to low-rated HMO	2.7 %	3.8 %	1.36	0.75	2.45
Assigned to low-rated HMO	N = 1614	N = 1679			
Stayed in HMO	76.0%	74.7%			
Switched to MediPass	14.1%	14.4%	1.03	0.75	1.39
Switched to high-rated HMO	9.9%	11.0%	1.13	0.79	1.60
Type II counties					
Assigned to high-rated HMO	N = 1087	N = 1037			
Stayed in HMO	70.5%	71.8%	-		
Switched to MediPass	29.5%	28.2%	0.92	0.68	1.24
Type III counties					
Assigned to low-rated HMO	N = 2097	N = 2153			
Stayed in HMO	76.3%	76.4%	-		
Switched to MediPass	23.7%	23.6%	0.99	0.79	1.23

[1]Consumer Assessment of Health Plans Study (CAHPS);

[2]Confidence Intervals;

[3]Health maintenance organisations (HMO);

[3]Medicaid primary care case management program (MediPass);

[4]High, low-rated: the reports or experience are composite scores that are averages of response to sets of individual items using four-category response options.

Farley 2002a

Europe PMC Funders Author Manuscripts

Table 3

Plan choices for April enrollees

	Mean or proportion	
	Reports (n = 2649)	**Control (n = 2568)**
Proportion choosing a plan	0.68	0.69

Farley 2002b

Table 4

Mean differences between actual and predicted monthly patient volume for the average outlier hospital, over 4 consecutive months in New York. After publication of a risk-adjusted outcome study, using ordinary least-squares regression*

State	Condition or procedure)	Outlier group§	Actual minus predicted monthly patient volume (95% confidence interval)			
			Month 1 (NY)	Month 2 (NY)	Month 3 (NY)	Month 4 (NY)
NY	CABG (target)	Better (D = 1.92)	13.4‡(4.3 to 22.6)	5.5 (−3.5 to −14.7)	6.7 (−1.5 to −15.0)	3.0 (−5.0 to 11.0)
		Worse (D = 1.91)	−4.0 (−9.0 to −1.0)	−7.1‡(−12.3 to −1.9)	−2.7 (−8.0 to −2.7)	−0.9 (−5.9 to 4.1)
NY	CABG-related (AMI)	Better (D = 1.96)	−4.9 (−12.3 to −2.4)	−1.4 (−8.7 to −5.9)	−1.9 (−8.7 to −4.8)	0.5 (−6.1 to 7.2)
		Worse (D = 1.38)	−4.5†(−8.5 to −0.6)	−1.2 (−5.2 to −2.8)	−1.6 (−5.4 to −2.2)	−6.0‡(−9.8 to −2.2)
NY	CABG related (PTCA)	Better (D = 2.14)	3.7 (−3.2 to −10.8)	1.1 (−6.0 to −8.3)	0.6 (−6.1 to −7.4)	−1.2 (−7.8 to 5.5)
		Worse (D = 1.34)	−2.6 (−7.0 to −1.8)	−1.4 (−6.0 to −3.1)	0.4 (−4.2 to −4.9)	−2.1 (−6.6 to 2.5)
NY	CABG-related (CHF)	Better (D = 1.74)	−2.8 (−8.7 to −3.1)	−4.0 (−9.9 to −2.0)	−0.5 (−6.0 to −5.0)	−1.7 (−7.1 to 3.7)
		Worse (D = 2.14)	−1.0 (−5.8 to −3.9)	−2.0 (−7.1 to −3.1)	−1.7 (−6.6 to −3.2)	−0.1 (−4.8 to 4.7)

*Positive numbers indicate that hospitals in that category had more admissions than predicted; negative numbers indicate that hospitals in that category had fewer admissions than predicted.

†Two-tailed P < 0.005

‡Two-tailed P < 0.01

§The Durbin-Watson statistics in this column represent the magnitude of autocorrelation affecting OLS models. Values close to 2 indicates the absence of autocorrelation.

Acute myocardial infarction (AMI); coronary artery bypass grafting (CABG); percutaneous transluminal coronary angioplasty (PTCA); congestive heart failure (CHF). Romano 2004

Table 5

Mean differences between actual and predicted monthly patient volume for the average outlier hospital in California, over 4 consecutive quarters after publication of a risk-adjusted outcome study, using autoregressive models (ARIMA)*

State	Condition or procedure	Outlier group§	Actual minus predicted monthly patient volume (95% confidence interval)			
			Quarter 1 (CA)	Quarter 2 (CA)	Quarter 3 (CA)	Quarter 4 (CA)
CA	AMI (target)	Better§	1.9 (−0.1 to −3.9)	−1.1 (−3.2 to −0.9)	−0.6 (− 2.7 to −1.6)	1.1 (−1.3 to 3.6)
		Worse§	0.7 (−1.6 to −3.0)	1.0 (−1.4 to −3.5)	0.0 (−2.3 to −2.4)	0.6 (−2.0 to 3.3)
CA	AMI-related	Better	−1.1 (−4.9 to −2.7)	4.2 (−0.1 to −8.5)	−3.8 (−0.8 to −8.3)	−0.1 (−4.6 to 4.5)
		Worse	1.0 (−1.5 to −3.6)	0.4 (−2.5 to −3.2)	0.4 (−3.2 to −2.5)	−1.0 (−4.1 to 2.2)
CA	Cervical discectomy (target)	Better	0.2 (−1.1 to −1.5)	−0.3 (−1.8 to −1.3)	−1.6† (0.0 to −3.2)	−0.6 (−2.2 to 1.0)
		Worse	−1.1† (−2.0 to −0.0)	0.3 (−0.9 to −1.6)	1.1 (−0.1 to −2.3)	0.9 (−0.4 to 2.1)
CA	Lumbar discectomy (target)	Better§	0.6 †(0.0 to −1.1)	0.3 (−0.3 to −0.9)	0.5 (−0.2 to −1.2)	0.8† (−0.1 to 1.5)
		Worse	−0.1 (−0.8 to −0.6)	−0.1 (−0.9 to −0.7)	−0.3 (−1.2 to −0.6)	−0.5 (−1.4 to 0.3)
CA	Discectomy-related	Better§	0.4 (−0.1 to −1.9)	−0.9 (−2.4 to −0.7)	−1.1 (−2.7 to −0.4)	0.4 (−1.4 to 2.1)
		Worse	−1.4† (−2.4 to −0.3)	0.2 (1.1 to −1.4)	0.0 (−1.2 to −1.2)	0.2 (1.0 to 1.5)

*Positive numbers indicate that hospitals in that category had more admissions than predicted; negative numbers indicate that hospitals in that category had fewer admissions than predicted. To estimate the total difference in patient volume for the average California hospital in each quarter, the numbers shown should be multiplied by 3.

†Two-tailed P < 0.005

‡Two-tailed P < 0.01

§The Durbin-Watson statistics in this column represent the magnitude of autocorrelation affecting OLS models. Values close to 2 indicates the absence of autocorrelation.

Acute myocardial infarction (AMI); coronary artery bypass grafting (CABG); percutaneous transluminal coronary angioplasty (PTCA); congestive heart failure (CHF) Romano 2004

Table 6

Mean changes in acute myocardial infarction (AMI) quality indicators in hospitals after publication of report cards for the early feedback group

	Early feedback (n = 42)			Delayed feedback (n = 39)				
	Baseline %	**Follow-up %**	**Absolute change % (95% CI)**	**Baseline %**	**Follow-up %**	**Absolute change % (95% CI)**	**Absolute difference for early versus delayed feedback % (CI)***	**P value**
All 12 AMI process-of-care indicators	57.4	65.5	8.2 (5.8 to 10.7)	56.5	63.6	7.1 (4.3 to 10.0)	1.5 (−2.2 to 5.1)	0.43
Process-of-care quality indicators:								
Left of standard admission orders	73.3	72.5	−0.8 (−5.9 to 4.3)	72.6	66.4	−6.2 (−13.7 to 1.2)	5.8 (−2.6 to 14.2)	0.17
Left ventricular function assessment	45.6	49.8	4.2 (−0.9 to 9.4)	39.3	46.9	7.6 (3.1 to 12.2)	−2.0 (−8.7 to 4.7)	0.56
Lipid test ≤ 24 h of arrival	34.1	51.1	17.0 (10.7 to 23.3)	35.7	54.9	19.2 (12.8 to 25.8)	−2.9 (−10.7 to 4.9)	0.46
Fibrinolytics ≤ 30 min or primary PCI ≤ 90 min	39.0	45.7	6.7 (−0.8 to 14.2)	35.9	43.1	7.2 (−0.5 to 15.1)	3.3 (−5.7 to 12.4)	0.47
Fibronolytic administration decided by emergency department physician	64.4	84.3	19.9 (10.7 to 29.1)	68.8	86.3	17.5 (9.2 to 25.9)	−1.6 (−9.5 to 6.4)	0.70
Fibrinolytics giver prior to transfer to CCU or ICU	80.4	95.7	16.3 (7.1 to 23.7)	85.5	91.9	6.4 (0.1 to 12.7)	5.8 (1.1 to 10.5)	0.02
Aspirin ≤ 6 h of arrival	75.9	82.6	6.7 (3.7 to 9.6)	72.8	77.1	4.3 (0.2 to 8.3)	4.3 (−0.1 to 8.8)	0.06
ß-blockers ≤ 12 h of arrival	28.3	73.7	45.4 (38.8 to 51.9)	32.2	71.3	39.1 (31.3 to 46.8)	3.1 (−5.8 to 12.1)	0.049
Aspirin at discharge	84.6	84.0	−0.6 (−4.2 to 2.7)	84.6	83.1	−1.5 (−6.5 to 3.4)	0.9 (−3.2 to 4.3)	0.75
ß-blockers at discharge	77.4	85.6	8.2 (5.4 to 11.1)	77.4	85.0	7.6 (4.1 to 11.2)	0.6 (−3.2 to 4.3)	0.75
ACE inhibitors or ARB for left ventricular dysfunction	75	81.7	6.7 (1.0 to 12.4)	71.6	77.0	5.4 (−0.8 to 11.5)	2.8 (−5.2 to 10.8)	0.48
Statin at discharge	57.6	85.5	27.9 (20.0 to 35.8)	57.8	85.8	28.0 (19.7 to 36.3)	−0.3 (−9.0 to 8.5)	0.95

Abbreviations: angiotensin-converting enzyme (ACE); angiotensin II receptor blocker (ARB); coronary care unit (CCU); confidence interval (CI); intensive care unit (ICU); percutaneous coronary intervention (PCI).

* Represents mean relative improvement in each indicator in the early feedback hospitals with the delayed feedback hospitals in the follow-up patient cohort after adjusting for indicator performance in the baseline patient cohort and type of hospital.

Positive values indicate better performance in the early feedback hospitals.

Tu 2009

Table 7

Mean changes in congestive heart failure (CHF) quality indicators among hospitals after publication of report cards for the early feedback group

	Early feedback (n = 42)			Delayed feedback (n = 39)				
	Baseline %	**Follow-up %**	**Absolute change % (95% CI)**	**Baseline %**	**Follow-up %**	**Absolute change % (95% CI)**	**Absolute difference for early versus delayed feedback % (CI)[a]**	**P value**
All 6 CHF process-of-care indicators	54.8	54.6	−0.2 (−5.0 to 4.6)	51.8	53.6	1.8 (−2.7 to 6.1)	0.6 (−4.5 to 5.7)	0.81
Individual CHF process-of-care quality indicators:								
Left ventricular function assessment	47.9	55.2	−7.3 (1.5 to 13.0)	43.4	52.5	9.1 (3.5 to 14.6)	1.2 (−5.3 to 7.7)	0.72
Daily weights recorded	14.8	24	9.2 (4.3 to 14.0)	15.1	22.7	7.6 (2.4 to 12.8)	1.8 (−5.2 to 8.8)	0.60
Counselling on ≥ 1 aspect of CHF	68.4	55.3	−13.0 (−21.8 to −4.5)	66.7	56.2	−10.5 (−18.2 to 2.7)	−0.4 (−8.4 to 7.6)	0.92
ACE inhibitor or ARB for left ventricular dysfunction	88.2	92.4	4.2 (0.7 to 7.8)	86.5	86.1	−0.4 (−7.4 to 6.5)	5.9 (1.0 to 10.7)	0.02
ß-blocker for left ventricular dysfunction*	40	71.7	31.7 (22.6 to 40.9)	38.3	67.7	29.4 (18.9 to 39.8)	3.5 (−6.1 to 13.1)	0.47
Warfarin for artrial fibrillation	52.4	64.2	11.8 (4.3 to 19.2)	49.3	63.6	14.3 (6.8 to 22.0)	−0.2 (−6.5 to 6.2)	0.96

Abbreviations: angiotensin-converting enzyme (ACE); angiotensin II receptor blocker (ARB); coronary care unit (CCU); confidence interval (CI); intensive care unit (ICU); percutaneous coronary intervention (PCI).

[a]Represents mean relative improvement in each indicator in the early feedback hospitals with the delayed feedback hospitals in the follow-up patient cohort after adjusting for indicator performance in the baseline patient cohort and type of hospital.

Positive values indicate better performance in the early feedback hospitals.

[b]Patients with documented ejection fraction of 40% or less during the index admission or within the previous 6 months were considered to have left ventricular dysfunction.