Comparative Effectiveness Research Review Disposition of Comments Report

Research Review Title: *5. Public Reporting as a Quality Improvement Strategy. Closing the Quality Gap: Revisiting the State of the Science.*

Draft review available for public comment from August 17, 2011 to September 17, 2011.

Research Review Citation: Totten AM, Wagner J, Tiwari A, O'Haire C, Griffin J, Walker M. 5. Public Reporting as a Quality Improvement Strategy. Closing the Quality Gap: Revisiting the State of the Science. Evidence Report/Technology Assessment No. 208. (Prepared by the Oregon Evidence-based Practice Center under Contract No. 290-2007-10057-I.) Rockville, MD: Agency for Healthcare Research and Quality. July 2012. Available at: www.effectivehealthcare.ahrq.gov/reports/final.cfm.

Comments to Research Review

The Effective Health Care (EHC) Program encourages the public to participate in the development of its research projects. Each comparative effectiveness research review is posted to the EHC Program Web site in draft form for public comment for a 4-week period. Comments can be submitted via the EHC Program Web site, mail or email. At the conclusion of the public comment period, authors use the commentators' submissions and comments to revise the draft comparative effectiveness research review.

Comments on draft reviews and the authors' responses to the comments are posted for public viewing on the EHC Program Web site approximately 3 months after the final research review is published. Comments are not edited for spelling, grammar, or other content errors. Each comment is listed with the name and affiliation of the commentator, if this information is provided. Commentators are not required to provide their names or affiliations in order to submit suggestions or comments.

The tables below include the responses by the authors of the review to each comment that was submitted for this draft review. The responses to comments in this disposition report are those of the authors, who are responsible for its contents, and do not necessarily represent the views of the Agency for Healthcare Research and Quality.

Source: http://effectivehealthcare.ahrq.gov/search-for-guides-reviews-and-reports/?pageaction=displayproduct&productID=1199
Published Online: July 26, 2012

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer 3	General Comments	I don't buy the premise – i.e., that public reporting is a QI intervention in its own right that can be evaluated the same way you'd evaluate a discrete clinical intervention. Public reporting is a vehicle for making provider-level performance information available to different constituencies (including consumers and the general public) for different purposes, but it's not a QI strategy or intervention in and of itself. What matters is how the information is used and to what end. To be sure, the argument that making performance information available to the public will lead to a more rational healthcare marketplace is one that has gained sway, but that's not the underlying (or only) rationale for measuring and reporting on performance. If the primary purpose of this review is to assess the evidence in support of this latter proposition, that's fine, but it should be framed in a way that makes this clear and acknowledges that this is not the only rationale for public reporting.	Thank you for your comment. This review is a part of a series of reports and quality improvement is related to the overarching goals of this series. The introduction has been revised to explain this and to acknowledge that public reporting can have other roles and goals.
Peer Reviewer 3	General Comments	I got the sense that the authors, too, struggled with this issue and had a hard time defining the scope of their systematic review. Although the title refers to "multiple pathways public reporting may influence quality of health care," they didn't provide a conceptual framework for understanding those pathways or the various mechanisms by which reporting performance data could foster quality improvement. Both target populations (providers and patients/purchasers) and key questions are defined in very broad terms, without reference to the context that might have helped better define the focus.	The text in the introduction has been revised to outline the different ways public reporting may influence quality. The context is included in our Analytic Framework because it is important to consider and text explaining this as well as the rationale for each key question has been added.
Peer Reviewer 1	General Comments	This review of the literature for the first time brings together a number of key questions and also the health care settings - all inclusive vs focused on one setting such as hospital.	Thank you.
Peer Reviewer 5	General Comments	It would be valuable to give some sense of timing of any effects. If there are no general statements that can be made, some information with the reporting of effects that speaks to how long or how rapidly changes were seen would be helpful.	This would be interesting to look at, however most of the research does not address this question. A few studies use lag variables to look at the impact of the public report at some point in the future but this is not done consistently.
Peer Reviewer 5	General Comments	Some of the writing implies causal effects. It would be valuable to avoid statement that report "X increased y", and instead use wording that shows associations of data to specific results.	We agree with your comment. We revised the text where we, as report authors, may have implied causality, and clarified when this was a statement by the article authors.
Peer Reviewer 5	General Comments	The organization of the paper that combines consumer/patient results with other results, such as provider or health plan, is confusing. It would be helpful to separate the consumer-specific data on its own.	Thank you for your comment. We have further clarified the population being studied: patient or provider. The stratification of the studies by setting and by key questions might help reduce any confusion.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer 5	General Comments	It was surprising that there were so few studies on consumer views, and these were mostly older studies. Given the explosion of the Internet and health information, it would seem likely that more recent research examined consumer beliefs and self-reported use of quality data.	We appreciate this comment. There might be studies available that explore consumer views (e.g. Angie's list) but they were not relevant in the context of this report. We were looking for studies that assessed the impact of public reporting of comparative quality. The studies that did not explore this relationship were excluded. As an additional check we selected studies we had identified on consumer views and conducted a citation search of these articles and reviewed the results. 90% of the articles identified in this verification search had been included in our search. We reviewed the remainder and this resulted in the inclusion of 1 new study.
Peer Reviewer 4	General Comments	This paper has some great value to add to the current understanding of the impact of public reports; however, certain key points needs to be clarified up front to set the context for the findings. To be meaningful to clinicians, report developers, the measurement community and others -- which I assume are the target audiences -- and to word the key questions clearly so that the findings are not misinterpreted, important points need to be made up front and certain concepts mentioned several times in the paper modified. Without such, the findings may be misinterpreted and applied in ways that dismiss the value of public reporting for consumers and purchasers in particular.	We appreciate this comment. We have provided more background and explanation when the key questions are first presented so that the result of this review do not give the impression that we dismiss the value of public reporting. We did not change the wording of the key questions as they were subject to TEP and series review and had been accepted when the draft was submitted for comment.
Peer Reviewer 4	General Comments	1) The actual content of a public report (e.g., measurement topics, unit of analysis) makes a significant difference in its usability for certain audiences and purposes. Throughout, this paper refers to assessments of public reports on "individual providers". Does this actually mean that the 29 studies were ONLY about reports that show scores of individual providers (e.g., a score for Dr. Jones rather than for Main Street Medical Clinic)? It is not clear from the narrative. Given the rarity of public reports that score individual providers, I assume that there wasn't a careful distinction made when selecting the studies used for this paper. Public reports that show scores for INDIVIDUAL providers and those that show scores only at the clinic or medical group level should not be grouped for evaluation of consumer use. Typically, the content of provider (non hospital, nursing home or home health) reports show results at the clinic or medical group level, providing information that is not useful for consumers to "choose a provider." This is an important point -- otherwise one might reach the conclusion that public reports on providers don't matter to consumers. They DO!	The results chapter on individual providers has been renamed "individual clinicians' to clarify. This section contains studies of public reporting in which the provider was actually an individual, mostly cardiac surgeons. The other chapters report on providers that are organizations (hospitals, health plans, long-term care providers). The one identified study of a medical group/outpatient clinic was about public reporting on fertility clinics. This was included with individual clinicians after consultation with AHRQ and the report editor as creating a chapter for one study did not seem useful.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer 4	General Comments	In addition, most reports on providers do not measure what consumers and purchasers think is useful -- cost, value, patient experience and overall outcome of care. These two limitations are common knowledge for the many groups who are engaged in consumer engagement and public reporting. The findings in this paper will have more credibility if this is explained up front and careful distinctions are made throughout so as to not reach conclusions that do not factor in these practical yet very important considerations.	Thank you for this comment. This issue was discussed in the Discussion/ Limitation section. A brief description, explaining this, has been added to the presentation of the key question that is partly meant to address this issue (Key Question 5).
Peer Reviewer 4	General Comments	In many areas of the country, the ability to choose among facilities or clinics does not exist, either because of being in a rural area or due to health plan network restrictions. Therefore, the purpose or impact of many public reports in the real world is about provider reputation and peer learning, and informing consumer decisions about their health and health care (e.g., importance of tests, following doctors' advice, etc). It is RARELY about choosing a different provider (see point #1 above). Again, folks involved in public reporting know this and nearly all have moved away from promoting public reports to choose a doctor because that frustrates consumers and impacts the credibility of the public report because the publicly reported scores are not at the individual provider level. (Many of us would love to see individual scores commonplace and public someday, but that is NOT the current state of reporting). This paper's repeated references to economic theory of perfect information to make consumer choices may have been the initial basis for public reporting, but the experience in creating the reports has led to different ways to impact and support improved outcomes for patients. Until we routinely have publicly reported scores for individual providers, this point about informing consumer choice only applies in limited cases of nursing home and hospital care (even then, for hospital care that is emergency or based on a doctor's referral or in a one-hospital region, the consumer isn't making a choice).	We share your thoughts on this topic. The introduction has been revised to clarify that public reporting can have purposes other than influencing patients' selection of providers. The economic as well as quality improvement theories are mentioned as background for the history of the initial motivation for public reporting as reflected in the literature. We have added mention of other influences.
Peer Reviewer 4	General Comments	This paper needs to clarify that it is an assessment of studies done on the current state of public reporting, which still has significant limitations regarding value for consumers and purchasers (see #1 and #2 above). And it was likely worse the further back in time one reaches. What was publicly reported 30 years ago and how it was done is different from what is being done today, so what is the methodological justification for grouping study results from the 1980's with research from the 1990's and to that completed in the past 10 years?	This is now stated in the introduction and also discussed as a limitation. We have shown trends in the number of studies over time and we have arranged results in chronological order where we felt it was important to elaborate on this very point.
Peer Reviewer 6	General Comments	This was a comprehensive undertaking and the project was well executed. The target populations are explicitly defined and the key questions are clear. Clinical meaningfulness is less relevant to this type of systematic review.	Thank you.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer 6	Clarity and Usability	The report is well structured and organized. Given the diversity of the available research, I appreciated that information was presented by qualitative versus quantitative studies.	Thank you.
Peer Reviewer 5	Executive Summary	Is HARM the correct terminology here? There are unintended consequences, such as a consumer declining to see a providing or leaving a health system. Yet does this constitute 'harm' in the conventional sense of medicine? We tend to think of harm accruing to individuals; this is an example of shifting care that may be of great benefit to a patient, but negatively affect the bottom line of a practice or hospital.	Thank you for the comment. The description of harms has been revised and it clarifies that harms includes negative consequences for patients as well as negative consequences for providers. The term 'harm' has been retained to maintain similarity across reports, but the text and definition explain that this includes negative, unintended consequences.
Peer Reviewer 5	Executive Summary	ES-8: define 'cream skimming' and 'cherry picking' ES9: 'not-Cardiac' – Does this mean care that is not cardiac care? Might help to explain initially. ES9, line 28: sentence starting with 'At issue…' is awkward. ES9, line 52: missing "the" ES9, last paragraph: wording is choppy and not clear ES14: KQ5 – 'mode and tone of message' – is this the quality data, or a communication to the consumer to review the quality data?	The definitions of these specific terms have been added to the text. The executive summary has been revised to match the revised text. We have attempted to clarify each classification and heading.
Peer Reviewer 5	Executive Summary - Methods (also applies to actual methods section)	The outcome is listed as improvements in quality of care, with "change in patient/proxy or purchaser health care behavior" as an intermediate goal. Change in quality is not necessarily the ultimate goal of public reporting. Transparency can lead to informed decision-making, which may or may not impact "quality of care". Consumer decisions can increase or decrease utilization of services, and may alter satisfaction in care or other consumer outcomes. Modifying consumer decisions are a product of being more engaged in their health or healthcare, but don't always lead to better health. It stands to reason, then, that the patient/consumer should also be identified as a final outcome target.	We appreciate this comment. However, within the context of the CQG series of reviews, improvement in health care is the outcome of interest even if this is not the only goal of public reporting. We have clarified this in the introduction. Because we realized that patient/consumer actions are important we included their changes in behavior as an intermediate goal, but we did not require that a study show the linkage to improved quality of care for it to be included.
Fred Edwards Public Comment		It was surprising to see that the public reporting initiatives of The Society of Thoracic surgeons was not mentioned. I have attached some excellent articles that appeared in Annals of Thoracic Surgery in 2011. You may find that they constitute a valuable contribution to the field. Thank you for your consideration.	Thank you for providing the article describing this initiative. Our reports synthesized research that evaluated public reports. We acknowledge that there are many more public reports available than are included. A limitation is that the review cannot analyze the impact of public reports that have not been studied, or if the study results are not available. An inventory of public reports would be different project.

Commentator & Affiliation	Section	Comment	Response
Jeff Maitland Public Comment		Approve without comments. On behalf of the American College of Chest Physicians (ACCP) the ACCP Quality Improvement Committee (QIC) appreciates the opportunity to comment on this report. The QIC would like to applaud the evidence based approach this document followed.	Thank you
Peer Reviewer 4	Introduction	Providers are motivated by reputational issues for sure, but also by the idea that purchasers (employers and health plans) may use the information for contracting which affects their income and network inclusion. This point needs to be added in several places where provider motivation is mentioned in the report. On page 3, 2nd paragraph, the stated contributions of this paper are too broad, particularly given the extremely important limitations noted in the general comments above. Consider refocusing?	We have added contracting as a motivation for providers' behaviors to the description particularly as studies of contracting behavior were included in the review. We revised the paragraph on contribution of the paper to be clear that this is the aspiration underlying the work. The limitations on the ability to accomplish this despite our efforts are included in the discussion and limitations sections.
Peer Reviewer 2	Introduction	Clear and Thorough	Thank you.
Peer Reviewer 3	Introduction	To better set the stage for this review, this section would benefit from a more nuanced discussion of the role of performance measurement in a complex adaptive system framework and of public reporting in that context – noting, for example, the differences (in content and level of detail) between process and outcome data that can inform internal quality improvement, more global measures of performance that might be used to assess quality from a policy perspective, and the kind of information that patients need to make decisions about their own, personal care in different situations.	More detail about the different roles of public reporting has been added to the discussion. Internal quality improvement is not addressed because it is outside the scope of this topic.
Peer Reviewer 3	Introduction	It would also help to describe how the locus and circumstances of decision-making vary by provider type, which has implications for consumers' potential use of the information in making health care decisions.	We have included this in the introduction and discussion.
Peer Reviewer 3	Introduction	This would also help develop more explicit research questions. I would suggest avoiding positing key questions in terms that presume simple causality – "Does public reporting RESULT IN improvements," "Does public reporting LEAD TO change" – since it's unlikely that any research would be able to establish such causality in this complex environment. Although I absolutely agree that differences in the purpose and context of public reporting are important, I didn't understand exactly what Key Questions 5 and 6 (page 4) were getting at, or what the difference was between them.	More detail has been added to the description of the key questions to clarify their meaning. We did not change the wording of the key questions as they were subject to the TEP and series review as well as peer review.
Peer Reviewer 6	Introduction	The background section of the report discusses the underlying assumptions of theories of economics and behavior change, but does not talk about the practical realities of transmitting information into the hands of users. Are these assumptions and theories still reasonable for public reporting of information related to quality of care? You might note their potential limitations, foreshadowing some of your conclusions.	Thank you for your comment. This has been added to the introduction.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer 6	Introduction	Either in the background or conclusions, you could discuss that alternative assumptions and/or conceptual frameworks should be considered. How does the existence of health insurance and existing patient-clinician relationships potentially interfere with the underlying assumptions? It is assumed that access to information about better health care options will induce individuals to change to health care providers, but is this realistic given that many individuals get their health insurance through their employer and have a limited of options from which to choose? They may also have established relationships with individual clinicians, and a lower than average quality rating for an organization may not be sufficient to persuade them to change.	The introduction has been expanded to describe the complexity of the decisions and the context for public reporting.
Peer Reviewer 1	Key Questions	Key Question #5 - another characteristic of public reporting program to consider might be whether the performance data is also linked to either QI tools for providers, or some health care recommendation for public.	Detail has been added to the description of Key Question 5 to indicate that these would be considered "characteristics of the report."
Peer Reviewer 1	Key Questions	Key Question #6 - contextual factor - what are other interventions that might impact the results - incentive programs, improvement collaboratives, campaigns, etc. Also what is the delay between timing of public reporting and the data collected - time lag issues, hypothesis being that the longer the time lag, the less impact once could see, data being less relevant and useful.	Detail has been added to clarify that the time lag and the timeliness of the data are characteristics of the report and included in key question 5. Other interventions fall under key question 6 and text stating this has been added.
Peer Reviewer 4	Methods	The third item in the criteria for quality assessment of individual included studies seems to be missing a word: "potential confounding (what?)".	More detail is provided in Appendix F and additional wording has been added to the text.
Peer Reviewer 2	Methods	Well-described Search strategy and appropriate	Thank you. Description of the search was revised to include the date the update was conducted while draft was being peer reviewed.
Peer Reviewer 3	Methods	In regard to Topic Nomination, it would help to explain, specifically, what the authors' charge was in undertaking this review, and how much latitude they may (or may not) have had in defining the focus	The paragraph on topic development in the Methods section has been revised to clarify this point.
Peer Reviewer 3	Methods	Regarding Search Strategy, the authors acknowledge that "public reporting does not map to standardized index terms in citation databases" (p. 6), but they do not explain the rationale for selecting the terms they used.	Wording has been revised to clarify how terms were developed and how searches were tested.
Peer Reviewer 3	Methods	Under Study Selection, it would help to know who the reviewers were (how many, what their background or training was), and whether they aimed for a certain level of inter-rater reliability before dividing the work among reviewers	These details have been added to the text.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer 3	Methods	Since they sought to be inclusive in their initial search, some of the exclusion criteria also seem puzzling and are not well explained – why rule out studies that focused on health care providers other than doctors and nurses (p.7, lines 23-24, p. 8, lines 36-37), for example, or studies that focused on quality data used for provider feedback, quality improvement, and benchmarking (p. 8, lines 3-4)? If the Population of interest is patients (or their representatives) and organizations that make health care or health care purchasing decisions, why include advocacy organizations (p. 9, lines 35-36) in the mix?	We appreciate this comment. Additional explanation has been added to the Study Selection section. Specifically, provider feedback and benchmarking do not include data that are public and were therefore excluded. Secondly, the population is patients, or anyone who might act for them, which could include advocacy groups who push for improvement in quality (though no studies were identified with this population). Lastly, the exclusion criteria were developed in consultation with experts to focus the review on general health care.
Peer Reviewer 3	Methods	The discussion of Outcomes needs to be tightened up, again, in the context of a framework that could help identify intermediate outcomes (including unanticipated adverse outcomes) from a dynamic systems perspective. In the Analytic Framework shown in Figure 1, I'd suggest that Harms (KQ2) should be shown as unintended outcomes (alongside intermediate outcomes on the right side of the figure), rather than as a side effect of the characteristics of reporting	The text has been revised to attempt to clarify the outcomes. The representation of the harms in an oval and with a curved line is a convention for a style of this type of figure that the authors have chosen to follow. It is considered a type of outcome. The graphic is not meant to imply that it results from the characteristics of public reports alone.
Peer Reviewer 3	Methods	Under Quality Assessment of Individual Included Studies, for reasons suggested earlier (namely, that public reporting takes place in a complex system environment), I question the appropriateness of criteria developed for the evaluation of controlled clinical trials – specifically, criteria 3 and 4 (p. 14, lines 51-53), which aim to account for "potential confounding" or "unintended exposure."	We appreciate your comment. However, we did not use criteria unique to randomized trials. We used criteria that have been suggested in the AHRQ methods guide as relevant to several types of observational studies as wellas RCTs. The only exception is the 1st criterion that asks if randomization was adequate for randomized trials or if the selection of the comparison group or time period is appropriate when studies were observational.
Peer Reviewer 3	Methods	Similar questions arise in regard to the use of criteria regarding "true treatment effects" to rate the quality of the studies (p. 15, lines 8-12).	The language is from the AHRQ methods guidance and is not meant to apply only to randomized trials. For this review treatment effect is interpreted as the impact of public reporting regardless of the study design.
Peer Reviewer 3	Methods	The rationale for not subjecting qualitative or lab-type studies to a systematic assessment (p. 15, lines 31-32) is also not explained	Thank you for this comment. Additional text has been added explaining that no guidance or prior practice dictated the choice of a tool for assessing qualitative work.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer 3	Methods	the method for Rating the Body of Evidence for Each Key Question (pp. 15-16)	The approach to rating the Body of Evidence follows the recommendations in the AHRQ Methods Guide. The text has been revised to clarify this and a citation is provided.
Peer Reviewer 3	Methods	I wasn't sure what the subheading Applicability (p. 16) referred to, although the narrative appeared to address questions regarding generalizability of study findings. It wasn't clear whether this referred to the studies under review or to the systematic review, itself. If the latter, it may be more appropriate to include this in the Limitations section at the end, rather than in Methods. If it relates to methods for the review itself, then it should be made clearer.	A definition and more explanation have been added to the text. Applicability is in a sense related to generalizability of the review findings which are dependent on populations and specifics of the identified and included studies.
Peer Reviewer 1	Methods	One comment to methods targets the issue of how the studies addressed the issue of confounding factors - how well or not they were able to consider other factors that might impact provider, organization performance - e.g. P4P programs, other incentives, national, regional campaigns or collaboratives that in addition to transparency would affect the results. This is complex - is there a way to assess the strength of studies in dealing with this - could be in the contextual factors?	How well a study addressed confounding factors was one of the six criteria used to assess the quality of the included studies. Given the nature of public reporting as an intervention, this criterion was given the most weight. A detailed description is provided in Appendix F.
Peer Reviewer 1	Methods	Also, how good is the concordance between the measures publicly reported and the area of improvement?	This is part of key question 5 and we have added text that explains that that it is considered, although it was rarely discussed in the studies we identified.
ACS-Public Comment	Methods	AHRQ used historical reviews and found limited sources to draw conclusions from in this draft report. …..the draft report used metrics which are not on par with current metrics specifically designed for public reporting. The initially reported metrics were used primarily for health plan or hospital accountability, resulting in limited clinical scope. It is important to highlight that we cannot draw generalizable conclusions without metrics that are specialty-specific. Therefore , the validity for the uses assessed in this manuscript will fall short of adequately answering the research question.	We appreciate this comment. This is discussed in the limitations of this review.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer 1	Methods/Future Research needs	One last comment here is that the report might discuss briefly the intent to update when new studies are published. I know already of one study - both quantitative and qualitative that looked at impact of public reporting on processes of care of physician practices and outcomes of care - submitted to peer review, not yet published, and there may be others - especially as this area of transparency is increasing over time.	As part of our evidence reviews, we have included a section on future research. For further details as to updating of reports within the EHC program, please see the chapter in the Methods Guide at http://www.effectivehealthcare.ahrq.gov/ehc/products/333/736/UpdatingCERs_MethodsGuideChapter_20110805.pdf. The Effective Health Care program through their established processes will assess proposed topics based on these criteria and determine whether a systematic review should be done or updated. There is not a standard time for updates in current policy
Peer Reviewer 6	Methods	The study team faced a significant challenge in that they were tasked with adapting methods recommended by AHRQ Effective Health Care Program to review a body of literature that was not focused on comparing the effectiveness of medical interventions. Given this challenge, the relatively long list of key questions, the variety of outcomes and settings, and heterogeneous nature of the literature, the study team did a commendable job.	We thank you for your appreciation of the complexity of this topic.
Peer Reviewer 6	Methods	In the quality review criteria, why did you not consider sample size?	That is an excellent observation and more detail is provided in Appendix F and additional wording has been added to clarify that sample size can be a consideration as part of the assessment of whether the comparisons are appropriate.
Peer Reviewer 6	Methods	For the type of data extracted from the studies, it would have been nice to include more information about the type of data in the public reports. I am not sure if there is a way to add information about this at this time - perhaps a paragraph?	We attempted to extract information about the format,content, and availability of the public report studied and this has been added to the text. However, this was rarely reported in the studies included in this review.
Peer Reviewer 6	Results	In p. 94, you refer to a mean of 5.56 out of 8 as “high” – perhaps this should be moderate to high	The text has been changed as suggested.
Peer Reviewer 6	Results	What do you see as the strengths and weaknesses of the measures (e.g., CAHPS, others) used in the studies? What are the implications of the measures used in interpreting the results? I suggest commenting on this somewhere in the Results section.	The discussion includes the importance of the measures used in the report, but an evaluation of these is beyond the scope of the review.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer 5	Results	1)line 15: changes in “process measures” or “patient experience” are very broad statements. What type? 2) line 20: surgery specific rates decline – for what procedure? Over what time? 3) line 45-47: consider change working to “designs prevented determining if…” rather than “impossible”. 4) Page 21, line 51-57: first sentence poor wording. Last sentence – wording suggests causal connection.	Thank you for this comment. We have made appropriate changes to the text to clarify these terms and issues.
Peer Reviewer 5	Results	1) line 26: high quality surgeon – what is this based on? Lower mortality rate? 2) line 32: what is ‘crowding out quality’? 3) line 40-43: wordy and unclear 4) line 46: add ? to end of sentence	The text has been revised to clarify these terms.
Peer Reviewer 5	Results	line 4: choices of younger patients: what choices were affected?	The text has been revised to clarify this point.
Peer Reviewer 5	Results	line 36: is employment a proxy for age? Line 36: influence selection of what?	Text has been revised to clarify this point.
Peer Reviewer 5	Results	1) line 3: add ‘s’ to finding 2) lines 10-12: remove statement about ‘not surprising, given…’ which is not a result but a discussion. 3) line 56: studies suggest harm….of what? (see prior comment about ‘harm’)	The text has been revised to clarify these points.
Peer Reviewer 5	Results	consider stating what CSRS is…hard to remember when there’s many abbreviations	We have expanded the abbreviations in more places to improve readability.
Peer Reviewer 5	Results	1) Line 27: patients were ‘lower risk’ – were they lower risk for surgery, or were there lower rates of surgery? Clarify. 2) line 34: change ‘then’ to than, and remove ‘was the most alarming result’ (take emotion out) 3) line 45: this paragraph has awkward wording 4) line 10: how is market share measured? 5) line 57: looked whether...insert “at”	The text has been revised to clarify these points.
Peer Reviewer 5	Results	Interviews & Surveys: might be valuable to categorize as older vs. more recent reviews. There seem to be movement from being against to tolerance to support (?)	The reporting on qualitative studies, which includes the interviews and descriptive surveys, has now been ordered by year of publication in all the results sections in order to allow identification of any trends.
Peer Reviewer 5	Results	Focus Groups: would group patient/consumer results and professional groups together, as opposed to citing all qualitative data together (consider subheadings separating them out)	Thank you for this comment. This change has been made in the results sections where they had not been organized in this way in the draft text.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer 5	Results	1) KQ4: impact on patients: be clear about patient choice vs. referral; information on consumer effect could be more explicit. Epstein seemed to describe referral patterns; is this patient or provider behavior? 2) Line 23-34: 'one reason..' is unclear how volume and market share is associated with consumer choice.	That is an important point. The text has been changed to clarify that the actors in the Esptein study are physicians acting on behalf of patients when they make referrals. Hence this is a type of selection (key question 4) and not about physicians who are the subject of public reports changing their behavior. The text has been revised to further clarify that volume and market share are used as measures of the numbers of patients selecting providers.
Peer Reviewer 5	Results	1) 1st para: who is saying volume change is associated with surgeons avoided? Author of paper, or EPC report authors? 2) KQ5: reference 33 – is this appropriately categorized as a report characteristic?	1) The text has been revised to clarify that this is the conclusion of the author of the paper. 2) Wording has been revised to clarify that this was the intent of the authors and older data could make report cards less useful unless it is predictive of current practice.
Peer Reviewer 5	Results	line 47: clarify 'crowding out quality in areas not measured'	The text has been revised to clarify the meaning of 'crowding out'.
Peer Reviewer 5	Results	line 52: clarify 'is not crowded out'	The text has been revised to clarify the meaning of 'crowding out'.
Peer Reviewer 5	Results	line 17: use of 'End Users' – be consistent with consumer terminology; consider 'consumer'	The heading has been changed to Consumers.
Peer Reviewer 5	Results	results are more detailed than prior results. First author – consider having consistency of details throughout report.	We have increased detail on other studies and reduced it in this section to improve consistency.
Peer Reviewer 4	Results	The charts comparing each study findings are an excellent resource.	Thank you.
Peer Reviewer 4	Results	In Table A of the executive summary and the conclusion (putting this note here as it's related to results): Given differences in reporting across the decades, add the date each study was conducted. Also, when it says "one report said..." I would like to see one of how many? Such as "Out of 7 reports, only one..." Or was there only ONE report that addressed that particular issue? It wasn't clear to me from the summary chart.	The total number of studies was added to the table and in the text to improve clarity.
Peer Reviewer 2	Results	Descriptions and Analyses of studies to be generally accurate. Disagrees with the importance of one of the studies and found one missing study. Hannan EL, Sarrazin MSV, Doran DR, Rosenthal GE. Provider Profiling and Quality Improvement Efforts in CABG Surgery: The Effects on Short-Term Mortality Among Medicare Beneficiaries, Medical Care, 2003;41(10):1164-1172.	We appreciate this information. The study has been added to the review.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer 2	Results	A relatively large amount of space is spent discussing the Apolito study (reference number 35) as evidence that mortality rates are higher in regions with public reporting and that harms occur as a result of public reporting. Given the very small size of the study, the fact that it refers to a subpopulation, and given that no information is provided in the study how many patients opted out of the registry in New York and non-New York hospitals, the results of this study seem very suspect to me. I am not arguing that the study should be removed from the report, but just that its importance is overstated.	Thank you for this comment. We agree and the text has been revised.
Peer Reviewer 2	Results	it is mentioned that reference 30 found that the number of NY CABG patients having surgery outside of the state declined. The context of the statement implies that this is yet another study that attests to the harms of public reporting, but in fact it shows that patients are undergoing CABG surgery in NY rather than being referred out-of-state. Either this statement should be omitted, or preferably, rephrased.	The statement has been rephrased to reflect the correct context.
Peer Reviewer 3	Results	My main suggestions regarding the Results section have to do with the organization of the material, which makes it very hard to follow	The organization of results section is now explained at the beginning for improved clarity.
Peer Reviewer 3	Results	I understand, and agree with, the rationale for presenting results by health care setting, especially when looking at impacts on consumers, since their degree of choice, timing of decisions, and autonomy of decision making vary considerably by setting. I also found the narrative descriptions of findings by setting to be very good, clear discussions, at an appropriate level of detail and with salient reviewer observations regarding the limitations or implications of the studies – for example (in the case of hospitals), the introductory section on Hospitals (p. 26), the Description of Quantitative Studies (pp. 28 ff), and the detailed analyses by key questions (pp. 31- 37). It would be helpful in these discussions to include information about the sponsors of the studies, where available.	Thank you for this comment. Sponsors have been included in the evidence tables. This has been added to methods and the overall discussion. The sponsors of research in this field are US government agencies and private foundations.
Peer Reviewer 3	Results	I don't understand the rationale for not integrating qualitative and lab-based studies into these discussions. Doing so would, I think, make for a much richer (and less fragmented) presentation of findings. The authors could acknowledge the limitations of the qualitative evidence (as they do the limitations of the quantitative studies). I would also recommend including the qualitative studies (identified as such) in Tables 3 – 7.	We appreciate this comment. However, the extent to which they could be integrated was limited by the lack of similarity in outcomes as well as the current state of art in doing this in systematic reviews. This is something we hope to continue to develop in future reviews.
Peer Reviewer 3	Results	It seems curious that survey-based studies were considered "qualitative," even when they were based on statistical sampling designs, with relatively large numbers (e.g., Schneider and Epstein, 1996).	Descriptive surveys were included with qualitative studies, whereas surveys that were used to collect data that were used in comparative analyses or modeling were included in quantitative studies.
Peer Reviewer 3	Results	In describing lab-based studies, it should also be noted whether the researchers used actual public-reported data (i.e., actual performance rates for actual providers) or mock reports.	Thank you for your comment. The methods section has been revised to clarify this.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer 3	Results	I also don't understand the rationale for separating studies of "cardiac" from "non-cardiac" reports on hospital care. At first, I thought this was intended to treat reports of cardiac surgery separately, since these data have been reported for longer periods of time in several venues, and since they represent the sort of elective procedures over which consumers have more control. However, findings in regard to the use of data about urgent or emergency heart conditions (heart attack and heart failure) would, it seems, be relevant to the use of reports about processes and outcomes of care for other conditions, such as pneumonia or surgical care. There may be a reason for separating these studies, but if so, it needs to be made clearer. Otherwise, I would recommend eliminating this distinction	We have added an explanation to the results for the rationale of separating cardiac and non-cardiac studies.
Peer Reviewer 3	Results	Where Summary results are retained at the beginning of each section (see comments below in section f), they need to be revised so that they can stand on their own. Specify, briefly, what outcomes were observed (and the direction of the outcomes) in what study populations, and include citations that identify these studies (not just the number of studies that supported one finding or another).	Thank you for this comment. More detail and study citations have been added to the Key Points at the beginning of each section; however they are not intended to stand alone, rather they are intended to introduce the main points.
Peer Reviewer 3	Results	The statements should be simple, objective statements of findings, with no editorial comments or speculation. For example, in the bulleted statement, "Employment status, likely a proxy for age, affected the likelihood that people would access comparative information about physicians (one study)" (p. 25, lines 36-37): (1) it is not clear whether the observation "likely a proxy for age" is from the study itself, or an observation of the reviewe (moreover, employment status may also be a proxy for insurance coverage); (2) it is not clearly stated in which direction the "likelihood" was affected; and (3) the particular study that had this finding is not indicated. Or, in the statement "Nursing homes that started with lower publicly reported quality ratings were more likely to improve their ratings than those that started with higher scores, which is not surprising . . ." (p. 26, lines 9-12), the latter part of the statement is an unnecessary editorial comment, more appropriate to the detailed narrative discussion. Also, avoid the use of undefined shorthand terms or jargon in these summary statements (e.g., "crowding out quality" [p. 22, lines 32-33], or "cream skimming" [p. 22, line 38]).	The text been revised to clarify when a statement is the assumption of the article authors. Furthermore, we removed editorial comments.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer 3	Results	Clarity and Usability: I found the organization of the Results section to be unnecessarily complex, confusing, and extremely difficult to follow without continually flipping back and forth to remind myself where I was. I'd suggest the following changes and reorganization: I agree with the authors that organizing the results by setting makes most sense (p. 20, lines 14-15) and would recommend eliminating entirely the Summary of Results by Key Question section (p. 20, line 35 through p. 26, line 12), which is unnecessarily repetitive of information that is summarized in later sections. Leave the Summary of Results by Key Question section as is, but eliminate the Overview of Findings sections in each of the subsequent sections of findings by Health Care Setting (i.e., p. 26 line 44 – p. 28, line 15; p. 57, line 34 – p. 58, line 24; p. 70, line 30 – p. 71, line 28; p. 85, line 44 – p. 87, line 8). These summaries are repetitive, and both are not needed.	A description of the organization of the results section has been added to clarify what is included and where there is repetition.
Peer Reviewer 3	Results	Revise summary bullets so that they are clear, objective statements of findings, complete with citations, as noted in comments in section d, above.	We have edited the bullets to clarify this point.
Peer Reviewer 3	Results	Be consistent when referring to Key Questions throughout the report. Identify them consistently by name (or abbreviated name) AND number.	We have edited the text to improve consistency.
Peer Reviewer 3	Results	Eliminate the breakdown between cardiac and non-cardiac hospital reporting (see earlier comments, in section d).	While we understand the basis for the comment, we chose to retain this organization for several reasons. Because of the amount of public reports pertaining to hospitals (both cardiac and non-cardiac), we decided to split them into the two sections - to present the findings in an organized, straightforward manner. Since approximately half of the hospital studies are about cardiac care we decided this division makes the most sense. Conclusions and the evaluation of the strength of evidence were done across settings, so they are combined in these steps.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer 3	Results	If possible, fold the discussion of qualitative studies and quantitative studies together (indicating in the narrative which are which, and what their limitations are), so that all studies addressing similar themes are reviewed together (see earlier comments, section d).	We appreciate your comment. We decided to report these separately for several reasons. Qualitative studies are not always included in systematic reviews. We wanted to include them but make it apparent to readers that they are different. Also the outcomes in most of the qualitative studies differ from the quantitative studies and do not exactly fit our analytic framework. Including them separately lets them serve as additional information without forcing studies that are extremely heterogeneous together.
Peer Reviewer 3	Results	Include qualitative studies, with appropriate references, in Tables 3 – 7.	Thank you for this comment. The qualitative studies are included in appendixes and not repeated in the text for space considerations. They are also not included with the quantitative studies because the information extracted differs and the qualitative studies were not quality rated.
Peer Reviewer 1	Results	In the sections of "Overview of Findings" (e.g page 21, 26, 57) it would be great to have the strength of evidence (fair, good, etc) and the reference for the studies that are summarized for each bullet. This information is in the tables, yet the reader is first exposed to the fact that there may be 2 studies showing positive impact, and 3 showing negative, etc - so at that point the reader would really want to know the strength of the evidence to gauge the significance of the differences, and have easy access to the citation to then go to the table and get the detailed information.	We revised our approach to strength of evidence and it is now across settings so we cannot add it to this section. However we have added it to the overview of results by key question. This comes before the other sections now.
Peer Reviewer 5	Discussion	1) para on Harms: any discussion about theories, validity, inconsistency with other findings? 2) starting line 33: discussion limited to change in providers. Assume this means not just individual clinicians, but also perhaps system-level changes in quality improvement, resources, feedback, education, monitoring…? Not explicated stated, and can certainly be the cause of any effects (as opposed to the public reporting, per se) 3) line 44: that fact that?	We appreciate your comment. The report has been updated to address these comments.
Peer Reviewer 1	Discussion	e. Discussion/ Conclusion: The discussion is fine. I might recommend some discussion about the myths that this review seem to chatter - e.g. that the impact on public reporting has been mostly on hospital QI, which is not really the case that comes out of this report. So pointing to some of the fallacies that are quoted broadly in the literature would be very helpful in anchoring the report even more strongly in its relevant findings.	The discussion now outlines broader impacts of public reporting.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer 1	Discussion	One area that might deserve more discussion is the fact that public reporting in itself may not have the scope of impact that are anticipated. But this is only ONE of multi-pronged strategies to improve care and performance. I wonder if there are studies that showed for e.g. that public reporting plus more targeted TA to low performers, or other tactic, might then lead to better results. This is complex for sure, yet I do think there needs to be an acknowledgment that PR is one strategy that combined with others may be effective.	We appreciate this comment. We have added that public reporting should be studied/viewed as part of a multi-pronged strategy.
(American College of Surgeons) ACS Public Comment	Discussion	In order to increase the potential for public reporting to be effective in influencing provider behavior, providers must value the reports as being reliable and valid…. To highlight the importance of this issue, ACS recommends that AHRQ includes importance of proper risk adjustment within the Discussion section of the report.	This has been added to the limitations in the Discussion Section.
ACS Public Comment	Discussion	Public reporting is still at early stages...These realities do not mean that better-directed and more efficient approaches will not be highly effective at improving care.	We have added text acknowledging that public reporting may be evolving and the review is limited to public reports that have been studied.
ACS Public Comment	Discussion	In the draft report AHRQ acknowledges that there are limited findings to determine whether public reporting is an effective intervention. The ACS recommends that AHRQ includes in their discussion that the level of measurement should be clearly conveyed to providers and patients so they can better understand the context of the measures. It is imperative to analyze the provider level of measurement such at the hospital level versus an ambulatory surgical center, versus an outpatient, versus a group level, and so forth.	While conclusions are drawn across different settings, the results are also presented separately, by hospital, health plan, individual clinician etc. We agree that more research in each area would allow more nuanced conclusions.
Peer Reviewer 1	Discussion /Future Research	Which brings me to the future research discussion. The authors could be more targeted in the recommendations about what research strategies would be recommended. That section outlines the challenges of research, yet does not provide enough of a path that funding agencies public and private, and researchers could begin to articulate and execute. The report is so rich in identifying what we know, what we do not know, where the gaps are - the report would gain great value if more specific recommendations would be presented about research questions, methods for example.	The future research discussion has been revised to improve usability.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer 4	Discussion/ Conclusion	See comments above, as they address limitations, findings and new research issues. In the limitations section, I disagree with the statement that public reporting is viewed as not generating new knowledge. Quite the opposite! What is the source for that generalization? Document it or delete it. Or clarify if you mean something else like knowledge based on original scientific or laboratory research.	We have revised the text to clarify. The comment is meant to convey that sometimes it is the research on public reporting that is not viewed as generating new knowledge that merits a journal article and therefore is not published. Our TEP members confirmed that evaluations of individual public reporting initiatives are important to the report producer but are not always published. If they were aggregated and synthesized they might provide insight, but this is difficult if they are not available.
Peer Reviewer 3	Discussion/ Conclusion	The discussion section would be stronger if it were framed in the larger context of current and future directions of public reporting -- for example, in support of value-based purchasing, administration-wide safety initiatives, and the proliferation of measures of performance and efficiency.	We have revised the discussion and added more detail. However we did not want to extend the comments beyond the report authors' expertise. We are hoping to work with experts in the field to build on the review to produce an article or other publication with broader recommendations.
Peer Reviewer 3	Discussion/ Conclusion	In the Limitations of the Review and Limitations of the Research on Public Reporting sections, the authors raise many of the questions and concerns that I had about this approach - including the appropriate use of assessment tools "rooted in the evaluation of clinical research," noting that there is "limited consensus about how to systematically assess evidence for questions in health services, public health, and quality improvement" (p. 109, lines 37-43). I agree entirely with their recommendation for the development of new approaches that integrate qualitative and quantitative evidence in systematic reviews (p. 110).	Acknowledged and included in Limitations and Future research.
Peer Reviewer 6	Discussion	Can you comment about why you believe the field evolved from looking initially at mortality data to the measures more commonly addressed at this time and the reasons for this evolution? With better risk adjustment measures now available, should mortality measures be revisited? Do you have suggestions for key measures to include in the public reports?	Thank you for this comment. Issues related to psychometrics, risk adjustment and the development or validity of the measures used to generate public reports were outside the scope of this review. For this reason the authors do not want to include opinions in the report that are not linked to literature we reviewed. This would be a good topic for another type of publication in the future.
Peer Reviewer 6	Discussion	I think it would be a good idea to mention the issue of health literacy in a report such as this given especially because individual consumers are one of the target audiences. What can be said about the various audiences' ability to process, understand and use the public reporting information when they do access it?	The text has been revised to include this idea.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer 6	Discussion	If people are not being exposed to all the public reporting data being developed with federal funding, what are your recommendations for future investment of research dollars?	Although we appreciate your comment, the report authors feel this type of commentary extends beyond the application of the results of the review and do not think this report is the best place for this discussion.
Peer Reviewer 5	Limitations	1) Page 108, line 52: add “but” after limits 2) Page 109, Limitation of Research: there should statements that talk about the limitations of the associations found in studies with positive effects. There could be changes in QM or provider behavior or consumer behavior in spite of the public reporting, or related to other activities occurring as a result of the reporting. Providers, health systems and employers could have intervened in other ways, in an effort that may not directly have to do with improving measures.	Confounding and the difficulty in studying/isolating the impact of public reporting are discussed in the report now.
ACS-Public Comment	Limitations	AHRQ acknowledges that there is limited evidence….The review indicates a great need for more controlled, prospective studies that can reduce bias.	The future research needs section attempts to acknowledge this.
ACS-Public Comment	Limitations	The IOM findings recommend that in order to have a successful public reporting program, reporting should be voluntary, confidential, and nonpunitive. The report also highlights that public reporting should make a distinction between quality and safety variables. ACS would argue that both of these important findings should be discussed in the AHRQ the report.	Thank you for this comment. By definition, confidential data provisions is not public and therefore was not included in this review. As the authors have not reviewed the literature on confidential feedback, they do not feel it is appropriate to attempt to discuss the relative merits of public reporting and confidential data provision.
ACS-Public Comment	Limitations	werecommend that AHRQ includes a note of caution within the report when analyzing data from public reporting interventions in other countries because results are based on different are economic, social, cultural and behavior forces. Few countries have the same biases as the US	We acknowledge that some differences across countries may be important but regional or socio economic differences in the US may be important as well. We provided information on the geographic location and the population included in all studies in the Evidence Tables and text. The findings from the studies conducted in other countries do not markedly differ from those in the USA and the number of non US studies is not enough to change conclusions.
ACS-Public Comment	Limitations of the Research on PR	ACS also recommends consensus development among stakeholders to agree upon appropriate methods for assessing the risk of bias, consistency, directness, and precision of the evidence for public reporting outcomes, with special focus on separating quality and safety for public reporting purposes.	The Effective Health Care Program of AHRQ develops methods for reviews. Drafts of methods guidance are posted for public comments and peer reviewed. This is an ongoing activity.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer 5	Future Research	While there is limited research on consumers, it is particularly interesting consumers were more engaged by patient satisfaction and patient recommendation data (as opposed to HEDIS and medical condition or procedure data). This speaks to greater need to understand what type of data can drive consumers to (a) access the data and review it, and (b) use it to gain knowledge and inform their decisions. This is an important area of inquiry that should be addressed.	We agree with this comment and the relevant text has been updated to reflect this.
ACS-Public Comment	Future Research	Bullet 1: ACS proposes that AHRQ includes a recommendation that future research should focus on supplementing Web materials and creating a national repository that can contain information on public reporting, while separating data on quality and safety performance. and CAPHS	We acknowledge the need for more information on public reports and we have included this in future research needs. We did not make specific recommendations on how this should be done because that is outside the expertise of the report authors.
ACS-Public Comment	Future Research	Bullet 2: AHRQ should consider the merits of the different levels of reporting and should recommend these levels of analysis in future reports. As a supplementary recommendation to increase public understanding of public reporting, ACS requests that AHRQ includes a recommendation for further investigation in which measures should be used use for accountability to a delivery system and which measures should be used to serve the public. We believe that there need for categorizing the reports is necessary to increase the ease of public consumption.	We agree that measures used for any reason, including public reporting, should match the purpose for which they are intended. Measures not used for public reporting are outside the scope of this review and the authors are not able to discuss these in this report. The commenter should consider proposing additional topics he believes merit further study to AHRQ through "Submit a Suggestion for Research": http://effectivehealthcare.ahrq.gov/index.cfm/submit-a-suggestion-for-research/
ACS-Public Comment	Future Research	Bullet 3: Therefore, a key policy research question for future studies is whether the extra cost and effort for widespread public reporting is worth the cost over timely, professionally maintained and shared information among relevant clinical groups.	Although interesting, the relative merits of public reporting verses confidential feedback to clinicians is outside the scope of this review.
ACS-Public Comment	Future Research	Bullet 4: Therefore, AHRQ should consider new and understudied methods of public reporting. Additionally, in the Future Research section, the report should recommend that the Department of Health and Human Services prioritize and provide support for the well-designed, mixed mode studies as well as pilot studies in future research, with additional investigation on the validity of administrative data.	We expanded the future research needs to incorporate the need for additional research including methods development to better address interventions like public reporting. Specific funding recommendations are outside the expertise of the report authors.
ACS-Public Comment	Future Research	Bullet 5: ACS recommends AHRQ further emphasizes the urgent need for investigation on how to best prevent perpetuation of disparities by highlighting this issue in the Future Research section.	Thank you for this comment. Disparities have been added to the text in the discussion.

Commentator & Affiliation	Section	Comment	Response
ACS-Public Comment	Future Research	Bullet 6: ACS would encourage AHRQ to look into alternative methods of obtaining individual-level physician data since the information on individual physicians on the Physician Compare Web site is relatively new and might lack validity. As such, we would recommend AHRQ to explicitly comment on the validity of the respective Compare Web sites individually as measurement tools.	Assessing alternative sources of data for public reporting is an interesting and important topic, but it is outside the scope of this review, as this review is focused on the evaluation of existing public reports.
ACS-Public Comment	Future Research	Bullet 7: the compelling need for revisiting all documentation requirements from regulations aimed at either quality, safety, and/or payment, which take tremendous time from care-givers on the front line and may actually detract from quality and safety of care through misleading information, fatigue, and hassle....ACS also recommends that AHRQ mentions that researchers consider the potential threat to validity as a result of new instruments used to collect information on outcomes.	We acknowledge the importance of the validity of measures and that there is a significant literature on this topic. However this aspect was outside the scope of this review.
Peer Reviewer 4	Methods:	Not sure where this comment fits: when referencing future research, the comment about linking interviews and observations to "administrative data" seems that it should also reference clinical or EHR data too. Why are studies in foreign languages potentially going to be included? Were they studies of public reporting in the USA? If not, it seems that cultural differences might be a complicating factor to justify exclusion from this study anyway.	While we did not identify a study that used EHR, we had added this possibility to the discussion. We did not exclude studies based solely on country or language because we believe there is information that can be learned from these experiences. Also it is a recommended standard for systematic reviews not to exclude by country or language alone. We found few foreign studies, but those we did find are included. We identified the geographic location of all the studies.
Peer Reviewer 5	Clarity and Usability	Clarity and Usability: The report is well structured and organized. Some of my comments speak to more separation of the consumer-related results. These are 'mixed' in with results of health plan, and sometimes, provider (qualitative results). Isolating the consumer-related data will strengthen the organization and flow. The main points are well presented. The summary Table 8 is very clear and nicely done. The conclusions point more to needed methodology strengthening of research and greater understanding of the behavioral and socio-technical aspects of using public data to gain improvement in care AND better patient informed decision-making.	Both consumers and provider are considered populations of interest in the review, but we have tried to separate the results when appropriate and clarify who we are talking about in the text.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer 4	Clarity and Usability	While the report is a wealth of information -- the charts in particular -- the narrative is too long if the intended audiences are people in the field engaged in measurement and public reporting. There is a lot of repetition of concepts which could be tightened up. This will be a major addition to the field of understanding public reporting, especially once the clarifications and issues noted above are addressed throughout the report.	We have attempted to do this within the required structure for these reports. The Executive Summary is designed as a shorter document that can stand alone and convey the key points.