Mechanisms and effects of public reporting of surgeon outcomes: a systematic review of the literature

Authors

Behrendt, K

Groene, O

Keywords

Surgeons; outcome assessment; Public reporting; Quality of Health Care; Quality Improvement; Task Performance and Analysis

Highlights

- The public reporting of individual surgeon outcomes in the England is unprecendeted in scope
- As yet, there is insufficient evidence of broad-scale public reporting of surgeon outcomes to be an incentive to improve quality
- Some studies suggest adverse patient selection after the introduction of public reporting of surgeon outcomes
- Public reporting of surgeon outcomes is often legitimized as empowering patients, but the data uptake in terms of discussions with patients is low
- Public reporting is most likely an incentive for low performing surgeons

- A number of key criteria need to be satisfied for public reporting of surgeon outcomes to be effective

Abstract

Background: Public reporting of surgeon outcomes has become a key strategy in the English NHS to ensure accountability and improve the quality of care. Much of the evidence that supported the design of the strategy originates from the USA. This report aims to assess how the evidence on public reporting could be harnessed for cross-country translation of this health system strategy; in particular, to gauge the expected results of the UK surgeon outcome initiative and to propose criteria that elucidate that prerequisites and factors that are needed to public reporting effective.

Methods: A systematic search of academic databases was followed by snowballing from the reference lists. Only peer-reviewed articles and primary studies were included.

Results: 25 studies from the USA (n=22) and the UK (n=3) were included. Suggestive evidence of a negative effect on access to surgery was found for high-risk patients and non-whites; one survey indicated presence of gaming. There was anecdotal evidence of quality improvement measures adopted by low-rated hospitals in New York. Most studies reported only on the effectiveness of public reporting, rather than addressing how effects accrue. This limits cross-country transferability of policy lessons. Based on our analysis, we propose factors impacting on the transferability of the evidence underlying the public reporting of surgeon outcomes, which may inform the adoption of this strategy in other health systems.

Conclusions: There is some evidence that public reporting can be an incentive for low performing surgeons to improve quality. Negative incentive on patient selection as suggested in the USA have not yet been observed in the UK.

Introduction

In the UK, an ambitious public reporting (PR) programme has been implemented in the last few years, steered by various high profile scandals about quality care outcomes [1,2]. In December 2012, NHS England published its 2013/14 planning guidance specifying - amongst other - the

aim to improve patient care through better commissioning of services, and to support patients in making informed choices, implemented via public reporting of (initially) ten specialities (better known as ‘Offer 2’) [3]. Today, PR is seen as a central tool for quality improvement and since 2013 individual surgeons’ outcomes are made public via the patient portal ‘NHS choices’ [4]. Data has been published for more than 5000 consultant surgeons in 12 specialties (Adult cardiac surgery, Bariatric surgery, Colorectal surgery, Endocrine and thyroid surgery, Head and neck cancer surgery, Interventional cardiology, Lung cancer, Neurosurgery, Orthopedic surgery and Upper gastro-intestinal surgery). Data sources and outcome measures vary among specialties, but all include mortality rates (MR) of patients.

The idea to publicly report provider outcomes in order to ensure and improve quality is not new. Nearly 30 years ago, the Healthcare Financing Authority (HCFA) published annual mortality rates of Medicare patients in hospitals in the USA in 1987. Shortly after that, in 1989, the most studied PR scheme, the New York Cardiac Surgery Reporting System (CSRS) was initiated. In New York State, hospitals needed to obtain a certificate in order to offer cardiac surgery, and mandatory PR outcomes have been enforced since 1991. From 1989 to 1992 mortality rates dropped from 3.5% to 2.8% (41% decrease of risk-adjusted in-hospital mortality) [5], while nationally it only decreased 18% between 1987 to 1990 (30-day MR). In the 1990s, Pennsylvania and Massachusetts also started surgery-focused PR.

Today, PR is a widespread tool to measure and advertise the quality of nursing homes, hospitals and healthcare-maintenance-organisations (HMOs) [6]. While in the English NHS surgeon outcomes were available before, they were not specifically released to the patients (but reported embedded on hospital web sites), not systematically reported in a common, standardized platform and not covering the wide range of specialities involved now. Thus, the recent publication of surgeon outcomes has generated substantial debate in academia and policy since its publication on NHS Choices in 2013/2014.

The assumption is that public reporting unfolds its effects via a range of pathways [7, 8]. Generally, a distinction is made between pathways referring to the *principle-agent-relationship* where patients are empowered by public reporting to make better choices, or pathways related to *competition theory*, whereas providers would improve outcomes to compete for patients. In practice, there are alternative mechanisms through which public reporting could affect outcomes. Werner [8] outlines 4 possible mechanisms: 1. patients might choose high-quality providers; 2.

GPs might refer to high-quality providers; 3. Purchasers select high-quality providers; 4. Providers themselves react in response to PR. In terms of patient choice, a growing body of evidence shows that PR of outcomes only have a limited impact: Patients have difficulties understanding comparative PR data, especially if different measures are contradicting [9, 10]. Other studies have found that consumers might value own experience or recommendation from family and friends more than comparative outcomes information [11]. In one study, fewer than 10% of patients looked at performance data [12]; in another, only 12% of patients who had undergone Coronary Artery Bypass Graft (CABG) surgery in NY knew about PR and fewer than 1% were able to state the performance of their surgeon correctly [13]. In terms of GPs referral to high quality providers, the proposition is that they would act as agents for their clients, but in fact evidence suggests that GPs also value informal sources of quality information more than performance ratings [14]. There is insufficient evidence on purchasing models for surgeon outcomes, even though this is an area of increasing importance as many health systems start to focus on purchasing for value rather than services. Research suggests that a large part of the impact of PR is amongst providers themselves [15-19].

There is a growing body of evidence about *what* the impacts of PR might be. Yet, a large volume of previous research on public disclosure focused at the hospital level, rather than individual surgeon level, and didn't sufficiently focus on the question by way of which mechanisms PR influences quality and which health system structures, governance arrangements and, broadly, prerequisites need to be in place for public reporting to be effective. However, this very question is a key to understanding the potential effects of public reporting and assessing whether research evidence stemming from one health system can be applied to another (here: the NHS in the UK).

The aims of this review are therefore to assess the evidence on whether and how public reporting of outcomes can be an incentive for surgeons to improve quality. Furthermore, we investigate whether the literature contains evidence regarding two key tenets of the UK surgeon-level outcomes public reporting initiative: (i) surgeons discussing their outcomes with patients, GPs, CCGs or managers in order to improve the quality of care (as indicated in the NHS planning document "Offer 2"), or (ii) public reporting leading to adverse patient selection, as feared by the surgical specialties [20]. We then (iii) discuss the findings and consider the potential impact of the current UK surgeon level reporting initiative and (iv) propose some key lessons and contextual factors that should be considered by other countries that are in the process of devising a surgeon level public disclosure strategy.

Methods

We gathered and assessed the literature on public reporting of surgeon outcomes, and synthesised the effects and mechanisms from this literature. Methods used to collect and analyse the information contributing to this review followed the PICO framework [21], PRISMA statement [22] and Hawker et al [23] tool for quality assessment of the retrieved papers.

Search strategy

Translating the question into search terms was informed by the PICO framework that has been recommended for literature review about clinical questions [21]. The framework includes the patient target group (P) , intervention (I), comparison group (C) and outcome (O). For the purpose of this review, a comparison group was not included into the search terms.

Target population were individual surgeons or surgery departments. To extend the number of potential hits, all known US states with a PR system or known USA-wide PR systems were included, as they might be mentioned in the title rather than “surgeon”. This approach was informed by the review by Totten et. al on a related topic [24]. In this way, all present literature, also from outside the USA, could be included.

As PR was the intervention, all synonyms for PR or outcomes were included in the search. As outcomes, all terms referring to consequences and assessment were included (see Appendix A).

Starting from these initial search terms, related Emtree and MESH terms were compiled and linked using Bolean operators: search terms were linked with OR and concepts with AND. The search was conducted on 15.07.2015 (Appendix B). The search was conducted in Medline, EconLit and Embase. The search was limited to all studies published after 1980, as the first PR initiative started in 1989. Further limits were human and English language. Due to the number of results, the search was further limited in Medline and Embase to title and abstracts. For the search in EconLit, only 3 concepts were applied and no limits, due to the much smaller number of results than in Medline and Embase.

Study selection

The initial results were screened according to their title and potential relevance to PR. In a second step the abstracts were considered. Of the remaining studies, the full text was retrieved if possible and pre-defined inclusion and exclusion criteria applied (Table 1). To make sure all relevant studies would be included, even if not found in the systematic search, reference lists of included studies were perused to find further relevant studies (snowballing). For the purpose of this review, only primary studies were considered. Literature reviews were excluded, but the reference lists of systematic reviews were used to further check that no relevant study was left out. To ensure the quality of the review, only peer-reviewed articles were included.

Table 1: Inclusion and exclusion criteria

Quality assessment

Cochrane Effective Practice and Organisation of Care Group (EPOC) provides detailed guidelines on how to assess the quality of studies, mostly clinical trials [25]. Due to the diverse nature of the studies included, the tool was found unfit to assess the quality of qualitative studies. To do justice to the different types of studies, a checklist developed by Hawker et al [23,26] was used, which can be applied to studies with qualitative and quantitative methods. It has 9 categories (abstract, introduction, methods, sampling, results, bias, analysis, transferability and implications) that can be rated with 1 to 4 points in their quality and added up to a cumulative score between 9 and 36. The results can be seen in Appendix C.

Data extraction

Again in accordance with Cochrane recommendations, the Cochrane data extraction worksheet [25] was adapted to fit the question and scope of this review. A meta-analysis of the study data was not attempted due to the various study designs and backgrounds.

Results

Studies found and evaluated

The database search led to initially 710 publications being found, of which 663 remained after exclusion of duplicates. After screening titles, 112 studies were identified as potentially relevant, of which 50 were chosen for full text reading after reviewing abstracts. 10 of these were included in the review. The screening of the reference list resulted into another 10 studies being located. Thus, in total, 25 studies were included into the review.

Figure 1: Flow chart of study selection process

Of the 25 studies identified, 8 had a survey or interview design, 8 an interrupted time series or difference-in differences approach, one case study, a case control Study, 5 cohort studies and 2 with a before-after design. 22 studies were from the USA and 3 from the UK. The period of analysis covered 1987 until 2012.

Key themes in the literature

The most frequent issue of PR addressed in the literature is the potential detrimental effect of PR on access to surgery – the assumption that surgeons might reject high risk patients because of the fear of becoming an outlier in the mortality statistics. All other themes were addressed far less commonly, particularly those related to the assumptions of “Offer 2”, namely that patients use the information to choose a surgeon or that GPs and purchasers use the reports to contract surgeons. Likewise, we didn´t find studies that describe the consequences of patients confronting the hospital/surgeon and requesting, because of their outcomes, a reassignment. Table 2 provides an overview on the studies included and on the outcomes of interest (Table 2).

Table 2: Included studies and outcomes of interest

Adverse patient selection was the most frequently studied outcome of public reports. Of the studies included in this review, 5 reported a positive or no effect on access to surgery, one a transient effect. The majority of the studies (n=14) however reported a negative effect on patient selection, suggesting that surgeons were less likely to accept high risk patients after the introduction of public reporting.

Table 3 summarizes the result of the studies on adverse selection as a consequence of public reporting of surgeon outcomes (Table 3).

Adverse selection as a potential consequence of public reporting

Of the studies addressing adverse selection, quality and strength of evidence varied considerably: For example, Hannan et al [27] reported a minor effect of PR, with 40% of NY cardiologists influenced by PR, but only 6% claiming to be strongly influenced. Narins et al [28] and Schneider et al [29] both reported a majority of cardiac surgeons being less likely to accept a high risk patient. Hannan et al [30] and Glance et al [31] found in a sample of 31 hospitals and 87 surgeons that high risk patients were less likely operated on by surgeons with high performance. This is consistent with the findings of Dranove et al [32], we found that the relative illness severity among AMI patients in NY and Pennsylvania declined, Moscucci et al [33] who detected a decline in the rates of comorbidities in AMI patients, while Omoigui found that patients referred from NY to Ohio were sicker than other referrals or patients from Ohio itself. Glance et al Hannan et al and Peterson et al on the other hand could not identify such effects in their studies [30, 31, 34].

Results from California are mixed: Li et al [35] and Romano et al [36] found no evidence of risk selection when they looked at patient case mix, but argue that high-mortality hosptals might have avoided high-risk patients. Moreover, Werner et al [37] found a transient negative effect on access to CABG surgery for African American patients, as they were less likely to receive CABG surgery. In the consecutive year (1996),

Mukamel et al [38] found that non-whites are more likely to receive surgery from low-quality surgeons. Overall, although study results are contradictory in details, there was some evidence that PR reduced access to surgery for very sick patients and patients from different ethnic backgrounds.

However, adverse selection is a topic that is difficult to assess, prone to a number of biases, and vulnerable to the availability of high quality data. Maythams et al study [39] suffered from potential responder bias due to low response rate of 50%, and, generally, the surveys conducted represent weaker evidence due to social acceptability bias. Narins et al [40] reported that 88% agreed or strongly agreed that "physicians may report higher risk conditions to improve their outcomes", while 57% of respondants in Brown´s et al study felt that surgeons and hospitals can manipulate the data [41]. Further, studies fail to clearly attribute exposure or provide control groups [42], suffer from missing data [43] or were conducted based on patient [44] or provider subgroups [45]. Moreover, even if adverse selection took place, this may not necessarily suggest lower quality of care: It has been suggested that the less invasive percutaneous coronary intervention (PCI) might be a substitute treatment for CABG patients being considered too sick to be offered surgery. Dranove et al [32] looked at PCI as a substitute for CABG in NY, where only CABG outcomes were reported. They also found that percutaneous transluminal coronary angioplasty (PTCA) procedure numbers decreased and that hospitals might have taken general measures to avoid high risk patients. Their dataset was, however, restricted to Medicare claims. Consequently, because of the diverse nature of study designs applied, systems and patient groups studied, the evidence of consequences of public reporting is limited and its impact on quality of care outcomes needs to be interpreted with great caution.

Table 3: Patient selection and access to care (continued on next page)

Table 4: Patient selection and access to care (continued)

Use of data by providers to improve quality of care

A key claim of “Offer 2” is that public reports could be used by the public to choose a provider, by GPs to inform referrals and by commissioners to contract for outcomes. This review found limited evidence for such interactions in health systems with a history of surgeon reporting. Two surveys assessed the usage of the data in interactions between surgeons and patients. Burack et al found that 29% discussed the dataless than weekly with their patients and 44% frequently (less than weekly) with colleagues [46]. Moreover, the study suggests that understanding of the methods underlying the published outcomes was limited amongst the surgeons. In the study by Hannan et al [47], 22% agreed that they routinely discussed the information in the report with their patients. Based on the literature, only a minority of surgeons seem to discuss their outcomes with patients on a regular basis. It has to be taken into account that all these studies were surveys from the USA; there was no data available yet from UK since the introduction of “Offer 2”. However, investigations and organizational changes amongst hospitals with suspicious surgeon outcomes could be a relevant implication of public reporting. Dziuban [48] and Chassin [49] interviewed staff from hospitals which were published as having high mortality rates. All hospitals were subjected to special measures and required to undergo a case review. In two cases a lack of dedicated cardiac surgery support staff was identified, the other two identified the approach to the treatment of emergency cases as the cause of the high mortality rate. Strikingly, Dziuban reported an increased team cooperation as one of the salient positive outcomes of the investigation into the high mortality rate. The reports demonstrate that the reasons for high MRs are complex and may not only be attributable to the surgeons themselves.

A proposition put forward by the proponents of public reporting is that learning about their performance compared to peers might lead to increased intrinsic motivation of surgeons to perform better as opposed to extrinsic motivation due to remuneration. We did not find research to support this proposition. Economic models exist to demonstrate the influence of intrinsic motivation on surgeon performance [50], however, such models rely on a large number of assumptions. A more pertinent implication of surgeon reports might be that low-performing surgeons cease practicing. Jha et al [51] and Hannans et al [47] studies indicate that a larger percentage of bottom quartile surgeons stopped practice in NY than top quartile. Similarly, in Hannans study, the decline in mortality rate was partly attributable to low-volume surgeons with high mortality rate that ceased to practice [47].

Discussion

The aim of this study was to assess the international evidence on public reporting with a special focus on the two contrasting theses surrounding the introduction of the UK surgeon-level outcomes public reporting initiative: improved quality through better reflection on outcomes by patients, GPs, CCGs or managers and adverse consequences of public reporting on patient selection

We found a substantial literature assessing potential and observed adverse consequences of public reporting and less literature addressing the mechanisms that translate public reporting into quality improvement initiatives, with the largest effects to be expected amongst the lowest performing providers. The findings need to be interpreted with great caution. The variety of methodological approaches (design with/without control group or surveys with risk of bias) make it difficult to reach a final conclusion on the presence and extent of adverse selection of patients according to risk [8]. Questions have also been raised concerning the reliability of underlying data sources, the extent and impact of missing data, the comparability of clinical and administrative sources, and variations in approaches to coding the data. Moreover, the mortality risk of a patient is influenced by the patient (e.g severity of illness or comorbidities) as well as the treatment. Since the beginning of PR, the appropriateness of the risk-adjustment procedures has been questioned and criticized. The NYCS has updated their risk-adjustment several times [52, 53] but research suggests that different methods of risk adjustments will lead to different results [54, 55].

More fundamentally, however, is the question to what extent the literature (mostly derived from the experience of public reporting of cardiac surgeon outcomes in NY state) supports a generalizability of the findings to the UK context. In the US healthcare system with its many health insurances and where surgeons are not necessarily employed by a hospital [56], PR might offer a stronger incentive for patient risk selection than in the UK, where surgeons are less dependent on the market. Thus, this evidence might not be applicable to the UK NHS (the only UK study about patient access to care could not found evidence of gaming). In order to assess the transferability of the resuls, the processes through which public reporting should lead to the desired effect will be elucidated in more detail below. Subsequently we will present some factors that policy makers may consider when devising a surgeon outcome initiative in their own country.

Whether public reporting can be an incentive for quality improvement can not be easily answered. From a behavioural economics perspective, PR data can be seen as a `nudge`[57], that means providing feedback to intrinsically motivated surgeons, who will then act accordingly and try to improve [58]. Kolstad´s study showed a higher impact f intrinsic motivation on NY surgeons' quality improvement than their revenue had [50]. In fact surgeons were willing to forego parts of their revenue to improve quality. An alternative behavioral economics mechanism for public reporting to drive quality improvement is via cognitive bias that can come into play leading individuals to fear deviating from accepted standards more than appreciating the potential to improve their care [59]. The prospect of being ´named & shamed´ might lead surgeons to change their choice of patients in order to avoid becoming an outlier, as discussed above. According to Kolstad, motivation is highest when expected and observed performance are far apart, whether better or worse than expected, yet the UK PR data as presented online mainly distinguishes between `OK´ performance and ´negative outlier´. As the media focuses on the negative outliers, this is a plausible mechanism for PR functioning. Chassin showed that low performers are more likely to leave the practice, so the PR might indeed have had an effect [49]. Transparency and consumer choice are also often mentioned as reasons to justify PR efforts [6]. Some doctors go as far as claiming that consent to a procedure can only be given, and be valid, if the patient is aware of the surgeon's performance [60]. Survey results from the USA suggest that mortality data is not often discussed between doctors and patients. This finding is in line with other research that shows that patients found it difficult to understand outcome measures and value more highly recommendations from family and friends [10]. In the UK, where patients traditionally have less choice, data usage might be even less significant and, should this be the case, PR a weaker incentive.

An assessment of whether public disclosure of surgeon outcomes can be translated to other settings also needs to consider the health care delivery system in which the surgeon operates. In this sense, public reporting can be considered as a complex improvement intervention of which the active ingredient is not well understood. For example, case studies of US hospitals that faced high MR show evidence that the published outcomes are rarely attributable to the surgeon alone. Organisational structures and a lack of dedicated staff and procedures can also have a large impact on patient outcomes [61]. Similarly, the three surgeons (out of more than 5000) who were found to be negative outliers in the UK so far faced an individual review as well as a review of their whole unit [62]. Organizational measures were taken rather than re-training of the surgeon [63-65]. On the other hand, a clear accountability can also be considered as a requirement to improve team efforts [66].

The absence of controlled studies means that is difficul to separate the effect of public reporting from other strategies. There is also evidence that non-PR feedback can improve quality as well: in New England outcomes feedback, hospital visits and training in quality improvement were combined. The mortality rate went down as various measures were taken at the hospital level– multiprofessional reviews on clinical processes, protocol implementation, reviews of deaths, hiring dedicated cardiac surgery staff, training and implementation of checklists [67]. This then also raises the question of whether an individual surgeon represents the right unit of analysis. In the USA, where surgeons might not be employed by the hospital, and in cardiac surgery where procedure volume is high, the answer might be different from in the UK, and for a different procedure (say, upper gastro-intestinal surgery, with its much lower caseload per surgeon). Whether the statistical power is sufficient to detect variations in quality outcomes is an issue we have raised previously, and a prerequisite for public reports to unfold their effects [68]. This depends mostly on the expected outcome, the volume of cases per surgeon and the number of surgeons performing the operation. These seemingly technical details are a good example to illustrate the limitations of transferring research evidence from one setting to another.

Ideally, in order to appraise research evidence to support translating a health system strategy from one setting to another one should a-priori consider the key factors that determine the translation. At the conceptual level (as reflected in the “generic quality criteria for cross-country comparisons of health systems and policies”), this implies appropriate use of theory, explicit selection of country comparator, rigour in the comparison design, attention to the complexity of the cross-national comparison, rigour of the research method and a clear contribution of knowledge for both theory development and policy learning [69]. In the contect of public reporting of surgeon outcomes, we have outlined in this review a wide range of the factors underlying the effectiveness of public reporting, a diverse set of theories to postulate the expected results and to formulate the implicit causal pathways, all of which will determine to what extent the published evidence is appropriate to justify the proposed health system strategy, in this case the formulation of “Offer 2”.

Based on the published literature and our analysis presented here, we suggest specific criteria for cross-country learning from public reporting initiatives, which may support other countries to assess the potential impact of public reporting of surgeon outcomes and to devising a public reporting strategy. These criteria cover health system factors, surgeon factors, data related factors, patient factors, and organizational factors (Textbox 1).

Textbox 1: Factors impacting on the transferability of the evidence underlying the public reporting of surgeon outcomes

Generic health system factors:

- Coverage and access: do all patients have equal access to the procedure?
- Choice: do patients have a choice of the provider (hospital, surgeon)
- Existing public reporting system: has some data (say hospital level outcomes) already been puiblished?
- What is the role of the GP in the referral system (gatekeeping, direct referral) and are there other means of patient referral or relevance in the context of public reporting?
- Purchasing mechanisms: how are services contracted for – based on volume or based on outcomes
- Number of surgeons conducting the procedure: how many surgeons care conducting the operation in the health system
- How will the results be published: is there a single entry portal?
- Does the single entry portal contain patient education material beyond the outcome data?

Surgeon factors:

- Employed or independent: are surgeons on a fixed-appointment with a hospital or clinic, or do they freely operate in different organizations
- Competition: do surgeons compete for patients?
- Type and volume of procedures covered: which procedure should be publicly reported? Depending on the case volume and accountability there will be major implications for the public reporting initiative

Data related factors:

- Is quality data available (clinical registry or high quality administrative data)?
- Have appropriate risk adjustment mechanisms for the procedure been developed?
- What is the median delay from procedure to registration to analysis to public report?

Patient factors:

- Does the entire population have easy access to the Internet?

- How is the health literacy level amongst the target population?
- Are there major regional variations in access surgery (limiting patient choice)?

Organizational and professional:

- Do hospitals/clinics have the capacity for team and hospital level discussions to learn from outlying performers?
- Is there a national agency in charge of liaising with low performers to ensure that appropriate investigations, where appropriate, are initiated?

Limitations

This review has several limitations. Due to the diverse range of studies and outcomes included, a meta-analysis of the could not be conducted. Some studies were not included due to lack of access, or due to not being published in a peer-reviewed journal. The studies included showed varying quality. No study addressed the the question of this review entirely, so that different pieces of information had to be put together and evaluated in relation to each other. Public reporting in the literature so far has mostly addressed mortality outcomes. The public reporting mechanisms will be the same for other outcomes, yet the effectiveness of mechanisms might differ. For example, patient reported outcomes might offer better metric properties (in the context of surgeons outocomes, where denominators are small) and might also lead for a more active engagement of patients [68]. Finally, most studies (n=22) found were from the USA with limited applicability to a UK context, therefore, we have emphasized the transferability of the results to different health system contexts..

Conclusions

PR represents a well established but not yet fully understood approach to improving quality of healthcare. As most research on PR is from a US context, the findings can't easily be transferred to the potential use by GPs and clinical commissioning groups in the NHS (or health care providers and purchasing agencies in different health systems). Therefore, if public reporting is introduced at population level, access to surgery and potential negative effects (adverse selection) should be monitored. Policy makers devising a strategy to publish surgeon outcomes

in their own health systems may benefit from contemplating the factors presented here that address the transferability of the evidence underlying the public reporting of surgeon outcomes and specifically, whether prevailing health system conditions are supportive of such a strategy.

References

[1] Marshall MN, Shekelle PG, Davies HTO, Smith PC. Public reporting on quality in the United States and the United Kingdom. Health Aff 2003;22:134–48. doi:DOI 10.1377/hlthaff.22.3.134.

[2] Department of Health. Learning from Bristol: The DH Response to the Report of the Public Inquiry into children's heart surgery at the Bristol Royal Infirmary 1984-1995, Executive Summary 2002:9. https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/273320/5363.pdf.

[3] NHS England. Putting Patients First - The NHS England business plan for 2013/14 – 2015/16 2013.

[4] NHS. NHS choices 2014. https://www.nhs.uk/service-search/performance/Consultants#view-the-data.

[5] Hannan EL, Kilburn H, Racz M, Shields E, Chassin MR. Improving the outcomes of coronary artery bypass surgery in New York State. JAMA 1994;271:761–6. doi:10.1001/jama.1994.03510340051033.

[6] Marshall MN. Dying to know: Public release of Information about Quality of Health Care. London: RAND, The Nuffield Trust; 2000.

[7] Berwick DM, James B, Coye MJ. Connections between quality measurement and improvement. Med Care 2003;41:I30–8. doi:10.1097/00005650-200301001-00004.

[8] Werner RM, Asch D a, R.M. W, D.A. A, Werner RM, Asch D a. The unintended consequences of publicly reporting quality information. JAMA 2005;293:1239–44. doi:DOI 10.1001/jama.293.10.1239.

[9] Zwijnenberg NC, Hendriks M, Damman OC, Bloemendal E, Wendel S, de Jong JD, et al. Understanding and using comparative healthcare information; the effect of the amount of information and consumer characteristics and skills. BMC Med Inform Decis Mak 2012;12:101. doi:10.1186/1472-6947-12-101.

[10] Damman OC, Hendriks M, Rademakers J, Delnoij DMJ, Groenewegen PP. How do healthcare consumers process and evaluate comparative healthcare information? A qualitative study using cognitive interviews. BMC Public Health 2009;9:423. doi:10.1186/1471-2458-9-423.

[11] Schauffler HH, Mordavsky JK. Consumer reports in health care: do they make a difference? Annu Rev Public Health 2001;22:69–89. doi:10.1146/annurev.publhealth.22.1.69.

[12] Dixon A, Robertson R, Appleby J, Burge P, Devlin N, Magee H. Patient Choice. Kings Fund 2010:215.

[13] Schneider EC, Epstein AM. Use of public performance reports: a survey of patients undergoing cardiac surgery. JAMA 1998;279:1638–42.

[14] Coulter A. Do patients want a choice and does it work? BMJ 2010;341:c4989. doi:10.1136/bmj.c4989.

[15] Marshall MN, Romano PS, Davies HTO. How do we maximize the impact of the public reporting of quality of care? Int J Qual Heal Care 2004;16 Suppl 1:i57–63. doi:10.1093/intqhc/mzh013.

[16] Williamson OE. The Economic Institutions of Capitalism. Free Press. New York: 1985. doi:10.2307/2392889.

[17] Holmstrom B, Milgrom P. Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design. J Law, Econ Organ 2001:7: 24–52. http://www.jstor.org/stable/764957?seq=1#page_scan_tab_contents (accessed July 2, 2015).

[18] Rosenthal MB, Fernandopulle R, Song HR, Landon B. Paying for quality: Providers' incentives for quality improvement. Health Aff 2004;23:127–41. doi:10.1377/hlthaff.23.2.127.

[19] Hanson K. Delivering health services: incentives and information in supply-side innovations. . In: R. D. Smith KH, editor. Heal. Syst. low-middle-income Ctries. an Econ. policy Perspect., Oxford: Oxford University Press; 2011, p. 103–24.

[20] Surgeons ask NHS England to rethink policy of publishing patients' death rates . The Guardian, 30.1.2015, https://www.theguardian.com/society/2015/jan/30/surgeons-nhs-england-patients-death-rates-bruce-keogh-jeremy-hunt-health [accessed 3.8.2016]

[21] Schardt C, Adams MB, Owens T, Keitz S, Fontelo P. Utilization of the PICO framework to improve searching PubMed for clinical questions. BMC Med Inform Decis Mak 2007;7:16. doi:10.1186/1472-6947-7-16.

[22] Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med 2009;6:e1000097. doi:10.1371/journal.pmed.1000097.

[23] Hawker S, Payne S, Kerr C, Hardey M, Powell J. Appraising the evidence: reviewing disparate data systematically. Qual Health Res 2002;12:1284–99. doi:10.1177/1049732302238251.

[24] Totten A, Wagner J, Tiwari A. Public Reporting as a Quality Improvement Strategy Closing the Quality Gap : Revisiting the State of the. Agency Healthc Res Qual 2012;Jul. Evide..

[25] Cochrane Effective Practice and Organisation of Care Group. EPOC-specific resources for review authors | Cochrane Effective Practice and Organisation of Care Group. Nor Knowl Cent Heal Serv 2015. http://epoc.cochrane.org/epoc-specific-resources-review-authors (accessed August 3, 2015).

[26] Groene O, Botje D, Suñol R, Lopez MA, Wagner C. A systematic review of instruments that assess the implementation of hospital quality management systems. Int J Qual Health Care 2013;25:525–41. doi:10.1093/intqhc/mzt058.

[27] Hannan EL, Stone CC, Biddle TL, DeBuono B a. Public release of cardiac surgery outcomes data in New York: what do New York state cardiologists think of it? Am Heart J 1997;134:1120–8.

[28] Narins CR, Dozier AM, Ling FS, Zareba W. The influence of public reporting of outcome data on medical decision making by physicians. Arch Intern Med 2005;165:83–7. doi:10.1001/archinte.165.1.83.

[29] Schneider EC, Epstein a M. Influence of cardiac-surgery performance reports on referral practices and access to care. A survey of cardiovascular specialists. N Engl J Med 1996;335:251–6. doi:10.1056/NEJM199607253350406.

[30] Hannan EL, Siu a L, Kumar D, Racz M, Pryor DB, Chassin MR. Assessment of coronary artery bypass graft surgery performance in New York. Is there a bias against taking high-risk patients? Med Care 1997;35:49–56.

[31] Glance LG, Dick A, Mukamel DB, Li Y, Osler TM. Are high-quality cardiac surgeons less likely to operate on high-risk patients compared to low-quality surgeons? Evidence from New York State. Health Serv Res 2008;43:300–12. doi:10.1111/j.1475-6773.2007.00753.x.[32] Dranove et al D, Dranove D, Kessler D, McClellan M, Satterthwaite M. Is More Information Better? The Effects of “Report Cards” on Health Care Providers. J Polit Econ 2002;111:555–88.

[33] Moscucci M, Eagle K a., 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. J Am Coll Cardiol 2005;45:1759–65. doi:10.1016/j.jacc.2005.01.055.

[34] Peterson E, DeLong E. The effects of New York's bypass surgery provider profiling on access to care and patient outcomes in the elderly. J Am Coll Cardiol 1998;32:993–9.

[35] Li Z, Carlisle DM, Marcin JP, Castellanos LR, Romano PS, Young JN, et al. Impact of public reporting on access to coronary artery bypass surgery: the California Outcomes Reporting Program. Ann Thorac Surg 2010;89:1131–8. doi:10.1016/j.athoracsur.2009.12.073.

[36] Romano P, Marcin JP, Dai JJ, Yang XD, Kravitz RL, Rocke DM, et al. Impact of public reporting of coronary artery bypass graft surgery performance data on market share, mortality, and patient selection. Med Care 2011;49:1118–25.

[37] Werner RM, Asch D a., Polsky D. Racial profiling: The unintended consequences of coronary artery bypass graft report cards. Circulation 2005;111:1257–63. doi:10.1161/01.CIR.0000157729.59754.09.

[38] Mukamel DB, Murthy AS, Weimer DL. Racial differences in access to high-quality cardiac surgeons. Am J Public Health 2000;90:1774–7.

[39] Maytham G, Kessaris N. A change in opinion on surgeon's performance indicators. Interact Cardiovasc Thorac Surg 2011;12:586–9. doi:10.1510/icvts.2010.257857.

[40] Narins CR, Dozier AM, Ling FS, Zareba W. The influence of public reporting of outcome data on medical decision making by physicians. Arch Intern Med 2005;165:83–7. doi:10.1001/archinte.165.1.83.

[41] Brown DL, Epstein AM, Schneider EC. Influence of cardiac surgeon report cards on patient referral by cardiologists in New York state after 20 years of public reporting. Circ Cardiovasc Qual Outcomes 2013;6:643–8.

[42] Bridgewater B, Grayson AD, Brooks N, Grotte G, Fabri BM, Au J, et al. Has the publication of cardiac surgery outcome data been associated with changes in practice in northwest England: an analysis of 25,730 patients undergoing CABG surgery under 30 surgeons over eight years. Heart 2007;93:744–8.

[43] Apolito R, Greenberg M, Menegus M, Lowe AM, Sleeper L, Goldberger MH, et al. Impact of the New York State Cardiac Surgery and Percutaneous Coronary Intervention Reporting System on the management of patients with acute myocardial infarction complicated by cardiogenic shock. Am Heart J 2008;155:267–73. doi:10.1016/j.ahj.2007.10.013.

[44] Joynt KE, Blumenthal DM, Orav EJ, Resnic FS, Jha AK. Association of Public Reporting for Percutaneous Coronary Intervention With Utilization With Acute Myocardial Infarction 2012;02115.

[45] Omoigui NA, Miller DP, Brown KJ, Annan K, Cosgrove D, Lytle B, et al. Outmigration for coronary bypass surgery in an era of public dissemination of clinical outcomes. Circulation 1996;93:27–33. doi:10.1161/01.CIR.93.1.27.

[45] Burack JH, Impellizzeri P, Homel P, Cunningham JNJ. Public reporting of surgical mortality: a survey of New York State cardiothoracic surgeons. Ann Thorac Surg 1999;68:1192–5.

[47] Hannan EL, Siu AL, Kumar D, Kilburn H, Chassin MR. The decline in coronary artery bypass graft surgery mortality in New York State. The role of surgeon volume. JAMA 1995;273:209–13. doi:10.1001/jama.1995.03520270043029.

[48] Dziuban SW, McIlduff JB, Miller SJ, Dal Col RH. How a New York cardiac surgery program uses outcomes data. Ann Thorac Surg 1994;58:1871–6. doi:10.1016/0003-4975(94)91730-2.

[49] Chassin MR. Achieving and sustaining improved quality: Lessons from New York state and cardiac surgery. Health Aff 2002;21:40–51. doi:10.1377/hlthaff.21.4.40.

[50] Kolstad JT. Information and Quality When Motivation Is Intrinsic: Evidence from Surgeon Report Cards. Am Econ Rev 2013;103:2875–910.

[51] Jha AK, Epstein AM, A.K. J, A.M. E, Jha AK, Epstein AM. The predictive accuracy of the New York State coronary artery bypass surgery report-card system. Health Aff 2006;25:844–55. doi:10.1377/hlthaff.25.3.844.

[52] Peterson ED. The need for “compassionate provider profiling”: refining risk assessment for percutaneous coronary intervention. J Am Coll Cardiol 2011;57:912–3. doi:10.1016/j.jacc.2010.10.022.

[53] Resnic FS, Normand S-LTLT, Piemonte TC, Shubrooks SJ, Zelevinsky K, Lovett A, et al. Improvement in mortality risk prediction after percutaneous coronary intervention through the addition of a compassionate use variable to the national cardiovascular data registry

cathpci dataset: A study from the massachusetts angioplasty registry. J Am Coll Cardiol 2011;57:904–11. doi:10.1016/j.jacc.2010.09.057.

[54] Shahian DM, Wolf RE, Iezzoni LI, Kirle L, Normand S-LT. Variability in the measurement of hospital-wide mortality rates. N Engl J Med 2010;363:2530–9. doi:10.1056/NEJMsa1006396.

[55] Grant SW, Grayson a D, Jackson M, Au J, Fabri BM, Grotte G, et al. Does the choice of risk-adjustment model influence the outcome of surgeon-specific mortality analysis? A retrospective analysis of 14,637 patients under 31 surgeons. Heart 2008;94:1044–9. doi:10.1136/hrt.2006.110478.

[56] Thaler R, Sunstein C. Nudge: Improving decisions about health, wealth, and happiness. Yale Univ Press 2008:53–71. doi:10.1007/s10602-008-9056-2.

[57] Bevan G, Fasolo B. Models of governance of public services: empirical and behavioural analysis of “econs” and ‘humans. In: Adam Oliver, editor. Behav. Public Policy, 2013, p. 530–1.

[58] Kahneman D, Tversky A. Prospect Theory: An Analysis of Decision under Risk. Econometrica 2007;47:263–92. doi:10.2307/1914185.

[59] Bevan G. Performance Measurement of “Knights” and “Knaves”: Differences in Approaches and Impacts in British Countries after Devolution. J Comp Policy Anal Res Pract 2010;12:33–56. doi:10.1080/13876980903076187.

[60] Clarke S, Oakley J. Informed Consent and Surgeons? Performance. J Med Philos 2004;29:11–35. doi:10.1076/jmep.29.1.11.30415.

[61] Pace N. Surgeons should beware of plans to print individual performance. BMJ 2002;324:979. doi:10.1136/bmj.324.7343.974.

[62] Taylor R. Surgeon performance data “misses the mark.” BBC News 2014. http://www.bbc.co.uk/news/health-30057602 (accessed October 21, 2015).

[63] Perry L. Just three surgeons named as having poor death rates as NHS publishes data for more than 5,000 doctors. Mail 2014;17.

[64] Ben James. Brighton heart surgeon named as having death rates above national average. Arg 2014. http://www.theargus.co.uk/news/11616753.Brighton_heart_surgeon_named_as_having_death_rates_above_national_average/ (accessed August 27, 2015).

[65] Donnelly L. Just three surgeons named as having high death rates. Telegraph 2014. http://www.telegraph.co.uk/news/nhs/11240241/PIC-AND-HOLD-Just-three-surgeons-named-as-having-high-death-rates.html (accessed August 27, 2015).

[66] Beed M, Brindley PG. Publication of surgical outcomes-data: whose team are we on? Br J Anaesth 2014;112:615–7. doi:10.1093/bja/aet432.

[67] O'Connor GT. A Regional Intervention to Improve the Hospital Mortality Associated With Coronary Artery Bypass Graft Surgery. JAMA J Am Med Assoc 1996;275:841. doi:10.1001/jama.1996.03530350023029.

[68] Walker J, Neuburger J, Groene O, Cromwell D, van der Meulen J. Public reporting of surgeon outcomes: low number of procedures leads to false complacency. The Lancet 2013; 382: 1674-7

[69] Cacace M, Ettelt S, Mays N, Nolte E. Assessing quality in cross-country comparisions of health systems and policies: towards a set of generic quality criteria. Health Policy 2013; 156-162

Figure 1: Flow chart of study selection process

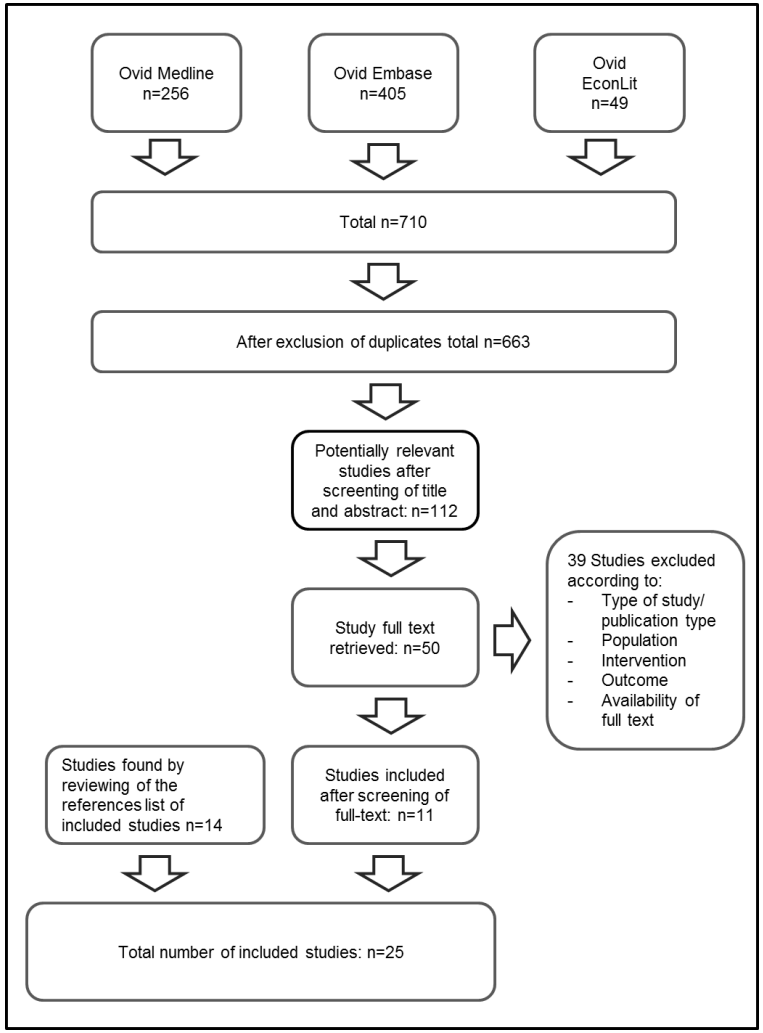
Ovid Medline n=256
Ovid Embase n=405
Ovid EconLit n=49
Total n=710
After exclusion of duplicates total n=663
Potentially relevant studies after screenting of title and abstract: n=112
Study full text retrieved: n=50
39 Studies excluded according to:
- Type of study/ publication type
- Population
- Intervention
- Outcome
- Availability of full text
Studies found by reviewing of the references list of included studies n=14
Studies included after screening of full-text: n=11
Total number of included studies: n=25

Table 5: Inclusion and exclusion criteria

Inclusion criteria	Exclusion criteria
- primary studies - all clinical areas - focus on supply side (healthcare providers) - main focus on surgery, either individual surgeons or surgery departments - mandatory PR or PR in general - information about possible consequences of PR, the *how?* - peer reviewed journals - full text available	- (systematic) Literature reviews - consumer ratings only (see definition of outcome) - focus on demand side - (duplication of findings in revision phase of manuscript preparation)

Table 6: Included studies and outcomes of interest

Studies included					Outcome of interest										
Name	**Year**	**Country**	**Identified in systematic review or snoballing**	**Study type**	**Patient selection**	**Complementary treatments**	**Gaming**	**Surgical training**	**Discuss data with patients**	**Discuss data with GPs and CCGs**	**Organisational changes**	**Intrinsic motivation**	**Caseload/ volume**	**Leaving profession**	**Quality score**
Apolito et al.	2008	USA, NY	sno	CC	x										34
Bridgewater et al.	2007	UK	sno	IST	x										35
Burack et al.	1999	USA, NY	sno	Q	x				x						31
Chassin et al.	2002	USA, NY	sys	Q							x				19
Dranove et al.	2003	USA, NY	sys	DID	x	x									28
Dziuban et al.	1994	USA,	sno	C							x				23

		NY													
Glance et al.	2007	USA;, NY	sno	C	x										35
Hannan et al.	1995	USA, NY	sno	C									x	x	34
Hannan et al.	1997	USA; NY	sno	C	x										34
Hannan et al.	1997	USA, NY	sys	Q	x				x						30
Jha et.al.	2006	USA, NY	sys	Q										x	25
Joynt et.a.	2012	USA, NY Mass Penn	sno	ITS/ DID	x										34
Khan et.al.	2007	UK	sys	ITS				x							31
Kolstad	2013	USA, Penn	sys	ITS								x			29
Li et al.	2009	USA, Cal	sno	BA	x										34
Maytham	2010	UK	sys	Q	x				(x)						29
Moscucci et al.	2005	USA, NY	sno	C	x										26
Mukamel	2006	USA, NY	sys	C	x										26
Narins et.at.	2005	USA; NY	sys	Q	x		x								29
Omoigui et.at.	1996	USA; Ohio	sno	ITS/ CBA	x										33
Peterson et. al.	1998	USA; NY	sno	ITS	x										27
Romano et al.	2011	USA, Cal	sno	BA	x										33
Schneider et.al.	1996	USA; Penn	sno	Q	x										32
Sherman et.al.	2013	USA	sys	Q	x			x							31
Werner et.al.	2005	USA; NY	sys	ITS/ DID	x										32

NY=New York; Penn=Pennsylvania; Mass=Massachusetts; sno= snowballing; sys=systematic review; ITS=Interrupted time series; Q=survey or interview; C=cohort study; CC=case-control study; DID=Difference-in-Difference Analysis; CBA= Controlled before-after study; CR= case report; BA=Before-after design

Table 7: Patient selection and access to care (continued on next page)

Author	Apolito et al.	Bridgewater et al.	Burack et al.	Dranove et al.	Glance et al.	Hannan, Siu, et al	Hannan, Stone et al.	Joynt	Kolstad	Li et al.
year	2008	2007	1999	2003	2007	1997	1997	2012	2013	2009
country	USA, NY	UK	USA, NY	USA, NY	USA, NY	USA, NY	USA, NY	USA, NY + Mass + Penn	USA, Penn	USA, Cal
study type	CC	ITS	Q	DID	C	C	Q	ITS/DID	ITS	BA
level	Pat	Hosp	Sur	Hosp, Pat	Pat	Sur	Sur	Pat	Sur	Pat
results	↓rates of coronary angiography for patients with shock and acute MI; ↑waiting times CABG. ↑ Mortality among NY patiens who were not revascularized	↑ numbers and percentages of low, high and very high risk patients significantly increased	70% reported no change in practice, but 64% claimed to have refused at least one CABG patient because of PR. More high risk patients were refused CABG surgery than surgery for aortic dissection, which is not publicly reported.	↓ relative illness severity among PR patients, as well as ↓ in the within-hospital heterogeneity of AMI patients. ↑ high risk patiens attending teaching hospitals. ↑Waiting times for CABG and PCI. Increased expenditures for AMI and CABG patients.	High risk CABG patients were found to be more likely to recieve surgery from a high quality surgeon	no sign of limited access to surgery or gaming	38% NY cardiologists stated the report had an influence on their referral practice, but only 6% said they were very much influenced by the reports	↓ likely to get PCI with acute MI, especially with ST-elevated MI or cardiogenic shock or cardiac arrest. ↔ overall mortality in reporting and non-reporting states. In Massachusets, odds of recieving PCI ↓ after PR had begun.	Komplex economic model, in which there was a minor effect of low-performing surgeons to avoid high-risk patients but this only accounted for 5% of the imporvement of quality.	27% less patients underwent CABG in 2006 than 2003; the case mix stayed unchanged however. Observed MR was lower in the high-risk quartiles in 2006 than 2003. All 4 outlier surgeons of 2006 met the expected MR in 2006.
effect	---	+	-	---	+	0	-	--	-	0

Table 8: Patient selection and access to care (continued)

Author	Maytham	Moscucci et al.	Mukamel et al.	Narins et.at.	Omoigui et.at.	Peterson et. al.	Romano et al.	Schneider et.al.	Sherman et.al.	Werner et.al.
year	2010	2005	2006	2005	1996	1998	2011	1996	2013	2005
country	UK	USA, NY	USA, NY	USA, NY	USA, Ohio	USA, NY	USA, Cal	USA, Penn	USA, IL	USA, NY
study type	Q	C	C	Q	ITS/CBA	IST	BA	Q	Q	ITS/DID
level	Sur	Pat	Pat	Sur	Pat	Pat	Pat	Sur	Sur	Pat
results	↓ surgeons claiming they will avoid high-risk patients, ↑claiming they will not change practice, ↑ thinking PR will improve outcomes	↓ patiens with AMI and cardiogenic shock underwent PCI, ↓ rates of comorbidities. Unadjusted MR ↓ but after adjustment the difference not significant any more.	whites were more likely to have access to low-RAMR surgeons. Level of access also dependent on health insurance type (for HMO patients worse than for FFS patients)	> 79% agreed that 1) knowledge of RAMR PR influenced decision to perform angioplasy, 2) patients might not get it due to PR RAMR; 3) RAMR PR influences decision on wether to intervene on patients with high expected mortality;	Patients referred from NY had a were sicker; ↑ average yearly volume of referrals from NY, while other referrals ↓, From 1989 MR among NY referrals ↑ than from other states or Ohio.	A smaller percentage of patients from NY recieved CABG surgery outside of NY.(from 12.5% to 11.3%). After PR was introduced, the odds of an older patient with AMI to recieve surgery increased significantly	high-mortality hospitals operated on slightly less high-risk patients (25% decrease in expected mortality)	59% of cardiologists found it more or much more difficult to found a surgeon to operate on a severly ill patient; 66% of cardiac surgeons were less or much less willing to operate on severly ill patients (compared to 3 years earlier	Majorities show concern with surgeons refusing high-risk patients, that high risk patients might be shifted to safety-net hospitals;	Before PR was introduced, white patients were more likely to recieve CABG than black patients. In the 9 years following the first PR, the racial disparities were wider. Then it got back to pre-PR level.
effect	-	--	-	---	---	+	-	-	-	+/-

NY=New York; Penn=Pennsylvania; Mass=Massachusetts; IL= Illinios;sno= snowballing; sys=systematic review; ITS=Interrupted time series; Q=survey or interview; C=cohort study; CC=case-control study; DID=Difference-in-Difference Analysis; CBA= Controlled before-after study; Case= case report; BA=before-after design; Pat=patient; Sur=surgeon; Hos=hospital; RAMR=risk-adjusted mortality rate

Appendix A: Search terms

Table 3: search terms according to PICO framework

P(atient)=target group	**I(intervention)**	**C(omparison)**	**O(utcome)**
surg* consultant* doctor* physiscian* OR known public reporting systems: New York hospital compare Medicare compare California Massachusetts Pennsylvania Cleveland Quality counts CAHPS HEDIS New Jersey Florida Vermont Virginia	ranking report card* public report* public disclosure score card* quality report* public release outcome* league table publication publishing mortality information service information dissemination access to information measurement health care quality quality control statistics scoring system outcome assessment mortality rate mandatory reporting public reporting performance measurement system benchmarking indicator compare score rank* rate metric evaluation	not included	quality perfom* incentive impact consequence result quality control health impact assessment program* impact motivation

Appendix B: Search strategy

	Embase Classic and Embase 15.7. 405 results
1	**consultant/ or surgeon/ or surgery/**
2	(New York adj3 Cardiac adj3 Report*) or hospital compare or medicare compare or California State Report Card or Massachusetts Health Quality or **(**Pennsylvania adj3 surg*)or (New York adj5 surg*) or Cleveland Health Quality Choice or (HCFA adj3 surg*) or QualityCounts or CAHPS or HEDIS or (new jersey adj 3 surg*) or (florida adj3 surg*) or (vermont adj3 surg*) or (virginia adj3 surg*)
3	1 or 2
4	information service/ or information dissemination/ or mandatory reporting/ or access to information/ or public reporting.mp. or public disclosure.mp. or public release.mp. or publication/ or publishing
5	Quality indicators, health care/ or Quality assurance, health care/ or Quality improvement/ or performance measurement system/ or health care quality/ or quality control/ or quality control/ or health care quality/ or report card* or scoring system/ or score card* or quality report* or outcome assessment/ or outcome*or league table*or mortality/ or mortality rate* or Quality Assurance.mp. or total quality management/ or performance measurement system/ or transparency or benchmarking or inform* or indicator*or compar*or score*or rank*or rat* or measure*or report*or release* or assess* or card*or metric* or evaluat*
6	3 and 4 and 5
7	incentive or impact or health impact assessment/ or program impact/ or consequence or result or motivation/or performance measurement system/
8	6 and 7
9	limits 1980-today, human, english
	Ovid MEDLINE(R) In-Process & Other Non-Indexed Citations, Ovid MEDLINE(R) Daily, Ovid MEDLINE(R) and Ovid OLDMEDLINE(R) 1946 to Present 15.7. 256 results
1	consultants/ or surgeons/ or physicians/ or General Surgery/
2	(New York adj3 Cardiac adj3 Report$) or hospital compare or medicare compare or California State Report Card or Massachusetts Health Quality or (Pennsylvania adj3 surg$)or (New York adj3 surg$) or Cleveland Health Quality Choice or (HCFA adj3 surg$) or QualityCounts or CAHPS or HEDIS or (new jersey adj3 surg$) or (florida adj3 surg$) or (vermont adj3 surg$) or (virginia adj3 surg$)
3	1 or 2
4	(public report$ or public disclos$ or public release or publication or publishing or information service or information dissemination or access to information).mp. or Consumer Participation/ or Public Participation/ or publications/ [mp=title, abstract, original title, name of substance word, subject heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier]
5	((((Outcome.mp. or Quality Assurance, Health Care/ or Health Care Quality Assessment/ or Health Care Quality Assurance.mp. or Healthcare Quality Assessment/ or Healthcare Quality Assurance/ or Quality Assessment, Health Care/ or Quality Assessment, Healthcare/ or Quality Assurance, Healthcare/ or Quality Control/ or Quality Improvement/ or Quality Indicators, Health Care/ or Quality Indicators/ or Quality of Healthcare/ or Task Performance.mp.) and Analysis/) or

	Task Performance/ or Task Performance, Analysis/ or Health Impact Assessment/ or Outcome.mp.) and Process Assessment/) or Assessment, Outcomes/ or Outcome Measures/ or Outcome Studies/ or Outcomes Assessment/ or Outcomes Research/ or ranking.mp. or report card$.mp. or score card$.mp. or quality report$.mp. or outcome$.mp. or league table$.mp. or mortality.mp. or measurement.mp. or quality control.mp. or scoring system$.mp. or mortality rate$.mp. or inform$.mp. or indicator$.mp. or compar$.mp. or score$.mp. or rank$.mp. or rat$.mp. or measure$.mp. or report$.mp. or release$.mp. or public$.mp. or publish$.mp. or assess$.mp. or card$.mp. or metric$.mp. or evaluat$.mp. [mp=title, abstract, original title, name of substance word, subject heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier]
6	3 and 4 and 5
7	motivation/ or incentive.mp. or impact.mp. or consequence.mp. or result.mp. or performance.mp. [mp=title, abstract, original title, name of substance word, subject heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier]
8	6 and 7
9	limit 8 to (english language and humans and yr="1980 -Current")
	EconLit 15.7. 49 results
1	(New York adj3 Cardiac adj3 Report$) or hospital compare or medicare compare or California State Report Card or Massachusetts Health Quality or (Pennsylvania adj3 surg$)or (New York adj3 surg$) or Cleveland Health Quality Choice or (HCFA adj3 surg$) or QualityCounts or CAHPS or HEDIS or (new jersey adj3 surg$) or (florida adj3 surg$) or (vermont adj3 surg$) or (virginia adj3 surg$)
2	(Consultant* or surge*).mp. [mp=heading words, abstract, title, country as subject]
3	(public report$ or public disclos$ or public release or publication or publishing or information service or information dissemination or (access to adj3 information)).mp. [mp=heading words, abstract, title, country as subject]
4	(Quality indicator* or Quality assurance or Quality improvement or health care quality or quality control or report card* or scoring system* or score card* or quality report* or outcome assessment* or outcome*or league table*or mortality or mortality rate* or Quality Assurance or transparency or benchmarking or inform* or indicator*or compar*or score*or rank*or rat* or measure*or report*or release* or assess* or card*or metric* or evaluat*).mp. [mp=heading words, abstract, title, country as subject]
5	2 and 3 and 4
6	1 or 5

Appendix C: Result of quality appraisal

Name		Abstract and title	Introduction and aims	Method and data	Sampling	Data analysis	Ethics and bias	Findings/results	Transferrability	Implications and usefulness	Total	Comment
Apolito et al.	2008	4	3	4	4	4	4	4	4	3	34	Further research areas not specified
Bridgewater et al.	2007	4	3	4	4	4	4	4	4	4	35	
Burack	1999	4	3	4	4	4	2	4	4	2	31	
Chassin et al.	2002	2	2	2	1	1	1	2	4	4	19	Only case study-part of the article taken into account
Dranove et al.	2003	3	4	4	2	4	2	3	4	2	28	
Dziuban et al.	1994	3	2	3	3	2	1	3	4	2	23	
Glance et al.	2007	4	4	4	4	4	4	4	4	3	35	
Hannan et al.	1997 a	4	3	4	4	4	2	4	4	2	31	
Hannan et al.	1997 b	4	3	4	4	2	2	3	4	4	30	
Jha et.al.	2006	2	3	2	4	3	1	3	3	4	25	only part about surgeons ceasing to practice considered
Joynt et.a.	2012	4	2	4	4	4	4	4	4	4	34	
Khan et.al.	2007	4	2	4	4	4	2	4	4	3	31	
Kolstad	2013	2	4	4	3	4	2	3	4	3	29	
Maytham et al.	2010	3	2	4	3	4	3	4	3	3	29	only 50% response rate, second survey has more participants
Moscucci et al.	2005	4	2	4	4	4	2	4	4	3	31	
Narins et.at.	2005	4	3	4	4	3	2	3	3	3	29	

Omoigui et.at.	1996	4	2	4	4	4	2	3	3	3	29	Limited generalisability since they only looked at one hospital
Peterson et. al.	1998	4	3	4	4	4	2	3	4	3	31	
Schneider et.al.	1996	4	2	3	4	3	1	3	4	3	27	statistical significance not reported for all measures
Sherman et.al.	2013	4	3	4	3	4	3	4	3	3	31	no exposure (not a PR state)
Werner et.al.	2005	4	3	4	4	4	2	4	4	3	32	