Public release of performance data and quality improvement: internal responses to external data by US health care providers

H T O Davies

Abstract
Health policy in many countries emphasises the public release of comparative data on clinical performance as one way of improving the quality of health care. Evidence to date is that it is health care providers (hospitals and the staff within them) that are most likely to respond to such data, yet little is known about how health care providers view and use these data. Case studies of six US hospitals were studied (two academic medical centres, two private not-for-profit medical centres, a group model health maintenance organisation hospital, and an inner city public provider "safety net" hospital) using semi-structured interviews followed by a broad thematic analysis located within an interpretive paradigm. Within these settings, 35 interviews were held with 31 individuals (chief executive officer, chief of staff, chief of cardiology, senior nurse, senior quality managers, and front line staff). The results showed that key stakeholders in these providers were often (but not always) antipathetic towards publicly released comparative data. Such data were seen as lacking in legitimacy and their meanings were disputed. Nonetheless, the public nature of these data did lead to some actions in response, more so when the data showed that local performance was poor. There was little integration between internal and external data systems. These findings suggest that the public release of comparative data may help to ensure that greater attention is paid to the quality agenda within health care providers, but greater efforts are needed both to develop internal systems of quality improvement and to integrate these more effectively with external data systems.

Keywords: quality of health care; quality improvement; comparative performance data; public disclosure

Background
Quality of care has risen up the health policy agenda in most developed nations over the past two decades or so. Significant quantitative data on clinical performance as one way of improving the quality of health care. Evidence to date is that it is health care providers (hospitals and the staff within them) that are most likely to respond to such data, yet little is known about how health care providers view and use these data. Case studies of six US hospitals were studied (two academic medical centres, two private not-for-profit medical centres, a group model health maintenance organisation hospital, and an inner city public provider "safety net" hospital) using semi-structured interviews followed by a broad thematic analysis located within an interpretive paradigm. Within these settings, 35 interviews were held with 31 individuals (chief executive officer, chief of staff, chief of cardiology, senior nurse, senior quality managers, and front line staff). The results showed that key stakeholders in these providers were often (but not always) antipathetic towards publicly released comparative data. Such data were seen as lacking in legitimacy and their meanings were disputed. Nonetheless, the public nature of these data did lead to some actions in response, more so when the data showed that local performance was poor. There was little integration between internal and external data systems. These findings suggest that the public release of comparative data may help to ensure that greater attention is paid to the quality agenda within health care providers, but greater efforts are needed both to develop internal systems of quality improvement and to integrate these more effectively with external data systems.

Keywords: quality of health care; quality improvement; comparative performance data; public disclosure

Department of Management, University of St Andrews, St Andrews, Fife KY16 9AL, UK
H T O Davies, reader in health care policy & management
Correspondence to: Dr H T O Davies
hd@st-and.ac.uk
Accepted 30 March 2001

www.qualityhealthcare.com
such data are made public, and how they respond when the data suggest that all is not well with their practice. In particular, the study sought to shed some light on how (or, indeed, whether) externally generated public reports on health care performance are integrated with internal strategies for identifying and dealing with quality problems.

**Approach**

The study used qualitative case study methods to explore attitudes to, and reactions to, externally driven comparisons of clinical performance. A qualitative approach was taken because of the desire to expose rich accounts of highly complex and contingent activities. Data gathering primarily involved qualitative semi-structured interviews with key stakeholders located in US health care providers, together with some documentary analysis of internal and external quality reports. The settings accessed, nature of the key informants, interview content, and analysis strategy are all explained below.

**SETTINGS**

Data gathering took place in six US hospitals all located in California. Purposive sampling was used to select centres with the reputation of being high quality providers renowned for the quality of their care. This strategy was used in an attempt to identify sites for fieldwork where there was likely to be more quality improvement activity to observe and explore—that is, the interest lay in examining leading edge centres rather than the middle majority or laggard. If the increasing emphasis on external data was bearing any fruit, then it is in these centres that there would be most to explore and learn.

Despite seeking centres with a high reputation, otherwise diverse institutions were included. Thus, two of the six centres selected were academic medical centres of international renown (indicated as Acad in the text), one was a hospital which was part of a group model Health Maintenance Organisation with salaried physicians (GM-HMO), two were private (but not for profit) medical centres (NFP), and one was a public provider “safety net” hospital (IP). This approach (seeking diverse settings) was taken to help buttress the external validity of the findings—that is, an exploration of provider responses in diverse settings should increase confidence that the findings were not case-specific. However, there was never an intention to make detailed comparisons between the individual case studies.

**INFORMANTS**

Within each setting interviews were sought with a range of key informants including the chief executive officer (CEO), chief of staff (CS; i.e. senior clinician with management responsibilities), senior quality managers (QM), chief of cardiology (CC; senior managing clinician in the cardiology service line), senior nurse manager (SNM), and two or three front line clinical staff (e.g. senior and junior physician and lead nurse within cardiology services).

With the exception of organisation-wide management leaders (CEO, CS, and QM), all informants were drawn from cardiology services. This service line was chosen for a number of reasons. Firstly, there exists a wealth of evidence about appropriate clinical practice in cardiology—for example, on the use of many categories of drugs. Secondly, there is ample evidence that actual practice often falls short of ideal practice in a number of areas—for example, in the use of low dose aspirin for patients at risk of myocardial infarction or in the timely use of thrombolytic drugs for those suffering from an infarct. Finally, there exists within cardiology both external systems of report cards—for example, the California Hospital Outcomes Project which reports public data on 30 day mortality after myocardial infarction—as well as confidential data systems designed for internal use—for example, a national register for myocardial infarction supported by Genentech, and the activities of a Health Care Financing Administration (HCFA) sponsored peer review organisation within the State.
ANALYSIS OF DATA
All tapes were reviewed immediately after each interview, with further written notes being prepared as necessary; the interviews were subsequently transcribed. The transcriptions were read through on several occasions by the author to highlight relevant data. Where necessary, the original tapes were replayed and contemporaneous notes were re-examined to clarify meanings and context. A broad thematic analysis,10 located within an interpretive paradigm,11 was used to identify and elaborate key themes. Statements relating to these themes were collated and cross checked to explore both strong themes and diversity within them. As the themes emerged, specific searches were made in the transcripts for countervailing arguments or beliefs and, where these occurred, they are reported. Cross case diversity was not explored.

Findings
The interviewees were first asked about their overall attitudes towards externally generated comparative performance measures, in particular their views when these data were made public. In the subsequent dialogue, informants were encouraged to reveal their perceptions about the strengths and weaknesses of such systems. Subsequently, interviewees were asked about the quality of care delivered in their own institutions, and were asked to describe the ways in which quality issues were identified and addressed. In particular, the interviewees discussed whether and how they reacted to external reports, and how these external data were integrated into internal quality improvement activities.

OVERALL ATTITUDES TO COMPARATIVE PERFORMANCE DATA
Attitudes to external comparative clinical performance data ranged from open hostility, through indifference and resignation, to reluctant acceptance and even guarded welcome. Negative comments included remarks such as: “I don’t think that data that are collated externally have had a positive impact—or any impact. I think they have had zero impact” (QM, NFP); and “they’re burdensome” (SNM, NFP). More grudging acknowledgements included “It’s a pain, but overall the care for the population improves . . . So that’s why I think that it [external monitoring] has to be there” (Physician, Acad); and “You get some benchmarks—trusted benchmarks” (QM, Acad), with even some enthusiastic support: “They’re welcome because we want to know how we compare . . . it helps us strive for improvement” (CS, NFP).

The range of these responses suggests, at best, ambivalence in welcoming the increasing use of comparative performance measures. Such ambivalence is seen within individuals as well as within organisations: “When it’s good news it’s ‘I love it, it’s great, this is me!’ If it’s not flattering, it’s like ‘Well, there’s something wrong [with the data]’” (CS, NFP).

Respondents who were more accepting of public scrutiny sometimes highlighted the fact that attitudes had shifted somewhat over the past decade—from hostility to greater acceptance—as the availability of comparative performance data had become commonplace: “No, I don’t think it bothers us now—we’re kind of used to it” (CC, Acad); “We’ve accepted the reality that it will be public and available” (CC, NFP); and “You just have so many people looking over your shoulder that that’s not troubling” (QM, PP).

Respondents were discriminating when welcoming or rejecting external review. For example, some were keen to differentiate between the potential benefits of confidential systems (such as the peer review organisations that provide comparative data within the State), and the much more problematic nature of public release of comparative data (such as the State mandated public release of health outcomes25 26).

CONCERNS ABOUT THE DATA
Compiling valid, reliable and meaningful comparative performance data is beset with pitfalls,14 15 22 and those interviewed were quick to raise a range of concerns. The essential fairness of the comparisons—and, in particular, the extent to which they took account of differences in patient populations or case mix—received considerable criticism: “A lot of the data is specious in that you can explain it away by patient selection etc” (CC, PP); and “Most of the time the data is [sic] not risk adjusted and the general population doesn’t understand what this means and so they take it at face value” (QM, NFP).

In addition, others highlighted the poor quality of the underlying data through, for example, inconsistent coding practices: “[This system] relies purely on administrative data, and administrative data’s so full of flaws” (QM, Acad); and “[these data are] generated by coders and medical records departments rather than by physicians themselves” (CC, Acad). Thus, apparent differences in performance were dismissed as artefacts of the data systems rather than seen as real clinical differences, and responses were often more concerned with reforming data collection and processing than addressing clinical care issues.

Finally, the long lags between data gathering and the production of official reports came in for considerable and scathing criticism: “Someone may ask you to respond to that information, but it’s so old that what we’re doing now has nothing to do with what was happening back then” (QM, NFP); and “It takes so long to develop the model and put the data through from all those organisations that, by the time we get it, it’s meaningless” (SNM, NFP). At the extreme, delays in the data reaching a public audience bordered on the farcical: “I was pretty astonished to read in a Sunday newspaper that [named unit] was considered probably the best in the city. I always felt it was very deserved. However the unit had closed three years before the article was written!” (CEO, NFP).

Withstanding the many negative comments on data quality, meaningfulness and time lines, there was a belief among some of the respondents that improvements were being seen—“over the years the data has gotten better in
terms of risk adjustment” (QM, NFP)—as well as a grudging acceptance that the deficiencies affected all providers similarly: “... it’s consistent, we all kinda use it in the same way and recognize its, uh...foibles” (SNM, NFP).

WHAT GETS MEASURED GETS ATTENTION
For all the accusations about the lack of meaning or relevance of the external data, many respondents expressed further concerns that, nonetheless, these data might distort clinical priorities: “We’re spending an awful lot of time and a very large amount of very finite resources to create a very elegant model [of post-MI mortality] that really looks at such a small part of what we should be concerned about” (QM, Acad). Thus, even before thoughts were turned to how external data might be used to improve care, study participants worried that “what gets measured gets attention”. Clinical issues highlighted by external data sets were thought to attract more institutional attention than was perhaps warranted—perhaps to the detriment of other unmonitored services: “There’s a different impetus when you know that the data has the potential to be released” (QM, Acad) and “It really fries people to do something to meet the task, rather than for clinically appropriate reasons” (QM, GM-HMO).

These concerns were not necessarily just academic. One provider reported that they had been pressured by an employers’ consortium purchasing group over some of the comparative data and had resisted what it saw as inappropriate priorities: “So we took the data back to [the purchaser] and said ‘That goal is not necessarily desirable. You’re pushing people to do something to meet the task, rather than for clinically appropriate reasons’” (QM, GM-HMO).

ACTING ON EXTERNAL PERFORMANCE DATA
Notwithstanding widespread concerns about the meaningfulness of external comparisons, providers do at times respond to the public release of comparative data. Given the importance they attach to public perceptions, this is perhaps unsurprising. Action seemed most likely when an organisation was seen to be performing poorly on any given external measure: “Being an outlier does motivate performance. There’s no doubt about that” (QM, GM-HMO); “Any time we do get really poor results, we will respond very um... very conscientiously” (QM, PP); and “Last time around we went from being the best to the worst in one fell swoop. It obviously got our attention more, shall we say, than if we had been the best” (Physician, GM-HMO).

Action to improve health care quality seemed rather less likely if data showed the organisation to be a “middle ranker”: “External indicators only have significance to us when we’re outside the norm—so we’ll tolerate middle of the pack” (CS, NFP); and “If you’re on the average it doesn’t give your hospital or your physicians
much of an incentive to look into the area—so that’s not terribly helpful” (QM, PP). However, there were also many instances cited where such complacency would not prevail: “I don’t think in the middle of the range is acceptable; we’re striving to be the best” (SNM, NFP); “If we’re in the middle of the pack it can be very upsetting” (CS, GM-HMO); and “whether we took action would depend on our own perception as to whether the data were an accurate reflection of what we think is happening” (QM, NFP).

However, a belief that actions could occur in the absence of an identified quality problem may be optimistic. When comparative data are largely exceptional, then these data tended not to be seen by front line workers but were filtered out by higher echelons within the organisation: “I wouldn’t even see it—unless it was bad” (Physician, PP).

RELATING EXTERNAL DATA TO INTERNAL QUALITY IMPROVEMENT

A strong theme to emerge from many interviews was that external data might “kick start” a process of internal enquiry, but that they were insufficient in and of themselves for complete understanding: “[External data] are the start of a process, you know, that really gets the ball rolling, in terms of an [internal CQI] investigation” (SNM, GM-HMO); and “We respond more to our own data, I think” (CC, Acad). So linkages between external data and internal quality improvement activities were generally weak. The weaknesses of these linkages arises from two distinct sources. Firstly, external data were generally found to be substantially out of date and thus lacking in relevance: “If you’re not doing it for yourself [collecting data] and reacting to it immediately, there’s a whole time lag and opportunities for improvement that you’ve missed” (CS, NFP), and “We definitely prefer in-house marking ventures) that provided the necessary practical resources for the analysis, presentation, and interpretation of such data—and a culture that encouraged, valued, and supported continuous quality improvement processes: “We have wonderful, wonderful motivated people, but if we didn’t have the resources to do this, we couldn’t. So there is resource. There’s not only people committed to excellence but there’s resources committed to excellence. That’s very important.” (CS, NFP); and “All the data in the world isn’t gonna help if the people at the top don’t wanna use it or don’t have the resources to use it” (CS, NFP).

In the absence of good local data and supportive resources, little quality improvement activity was seen: “We don’t do it [benchmarking] and we don’t have the resources to do it... really, no way, since we don’t have ongoing databases” (QM, PP).

ENCOURAGING SERVICE DEVELOPMENT AND PRACTITIONER CHANGE

In none of the organisations were significant financial incentives used as levers for change. More commonly commented upon was the fact that reward structures were sometimes disincentives to high quality—for example, salaried physicians attracting additional workload as a consequence of a reputation for excellence or fee-for-service reimbursement encouraging throughput over excellence: “The major emphasis is on access and throughput... I think that outcomes are secondary” (QM, PP).

Although better alignment of physician rewards was thought sensible, few respondents were interested in using financial incentives to drive practitioner change. Instead, the key issues for pressuring change were seen as credible comparative data of quality problems and detailed exploration of clinical processes, coupled with professional and institutional pride. “So I do see physicians taking it very seriously, they do want that data to reflect favourably on them, there’s a tremendous pride in their work” (CS, NFP) and “If you are sort of an outlier, that’s going to, without anybody saying anything, influence your behaviour” (Physician, GM-HMO). Thus, identifying and dealing with quality issues were seen as indicators of peer esteem and good professional practice: “If you’ve got the best outcomes [and] least complications, you have a higher standing with your peers. And if you know you’ve got a problem and you address it, that improves your standing... They [physicians] are also very competitive. They want to do the right thing, and they want to do it as well or better than everybody else” (CS, NFP).

The greater openness fostered by the report card movement—in itself legitimising a greater openness within institutions—was thus seen as a very important means of encouraging more reflective practice. The availability of good comparative data can then work to enhance and channel intrinsic motivations: “Physicians are self-correcting, they’re very competitive, they always want to be the best. If you show them data and they’re not as good as their partner, they tend to try and figure out themselves what’s going on... We’ve been trying to use it [comparative data] in a non-punitive, self-correcting mode” (QM, NFP).
Conclusions
The public release of comparative performance data has grown to greater prominence in health care in many countries. Public policy and considerable private sector activity have both contributed to these trends, but relatively little research is available to shed much light on whether and how such a strategy might improve health care quality. Indeed, although many rationales are available and have been articulated, current schemes tend to be vague about the purported mechanisms of action whereby public release will improve health care quality.12

This study sought to get inside health care provider organisations to explore the dynamics as they respond to more public scrutiny of what have hitherto been confidential professional matters. It is because current best evidence suggests that health care providers should be the key targets for publicly released comparative performance data12 that it is important to understand the mechanisms by which such data might be actioned.

The key findings from these interviews can be summarised as follows:

- The growing availability of comparative performance data, both internally and externally driven systems, have made quality of care issues much more visible than hitherto, hoisting them higher up the providers’ agenda.
- External data systems turn up the heat on health care provider—most especially so when these data are made public—and encourage them to examine the clinical issues covered by the measures.
- The accuracy, validity, and timeliness of external data sets are widely called into question, severely limiting their legitimacy in the eyes of health care providers.
- Despite perceptions about the inadequacy of the measures, many providers are worried that “what gets measured gets attention” and thus raise fears that disproportionate attention may be paid to those clinical areas on which data are publicly released.
- External data have greatest impact when they indicate that performance is below that expected. For some providers, anything less than exemplary performance creates a desire for action. For many others, however, so long as the external data do not indicate that they are significantly worse than average, no actions would result.
- Wherever possible, providers seek verification of any problems identified from outside by reference to internal data sets and subjective assessments based on “soft” data.33 Internal data sets tend to cover clinical processes in considerable detail, in contrast to external systems which often focus on health outcomes.
- Peer pressure, professional pride, and the relentless logic of credible comparative data were seen as the key drivers of changes in individual behaviour rather than financial or other external incentives.
- The public release of comparative data offers one way of building pressure on health care providers to prioritise health care quality issues.

Nonetheless, there is a considerable way to go before these data will be seen as both timely and credible when they appear to criticise local practice. In practice, attempting to win over providers with more credible data, or attempting to shorten the data delivery time to one that is acceptable, may be difficult—and may not even be necessary. This study suggests that the data just need to be credible enough to prompt further local investigations. What is clear is that effective local quality improvement activity is predicated on the availability of detailed process based clinical information and the resources to enable the exploration of this. Yet currently there is little connection (never mind integration) between internal and external data systems. This would seem to be a lost opportunity. The growing availability of voluntary based, bottom up, clinically driven comparative data bases—which emphasise a combined analysis of both process and outcomes—may offer some potential to bridge this gap.34

Caution should be exercised in extrapolating from this analysis to other nations or contexts.35 A study of this type has a number of important limitations. Most obviously, the study took place in California at a time when health care providers are under considerable pressures to cut costs as aggressive managed care begins to bite. Nonetheless, most health care providers in developed countries are familiar with stringent financial circumstances. In addition, the accounts presented reflect only the perceptions and conscious constructions of the stakeholders interviewed. Only very limited corroborations of the accounts was sought—for example, through sight of quality improvement reports or through cross referencing between interviews in the same organisation. The potential certainly exists for these accounts to be inaccurate or incomplete. Nonetheless, all the participants were willing volunteers for the study (there were no significant refusals) and gave every sign of being engaged and thoughtful with the subject. The academic nature of the study and the independence of the interviewer also contributed to a spirit of open enquiry.

Despite these notes of caution, there are of course many similarities across health care providers even in different countries. The public release of comparative performance data is an international phenomenon, and commonality of experience in responding to these data may be as important as diversity. Thus, the findings from this study should stimulate debate about the appropriate development of comparative data systems in many countries and settings.

The public release of comparative clinical performance data has become a “de facto” health policy in most developed nations. Whereas previous debates have largely revolved around the technical issues of data collection, analysis and interpretation,36 we now need to be much more concerned with how such data are used—for good or ill—within health systems. For example, it is still far from clear that any benefits arising from the public release
of comparative data will outweigh both the costs and the harms incurred. Since improved clinical processes within health care provider organisations will be the main way that real improvements are delivered, it is here that we must seek the evidence. It is here too that we need a better understanding of the dynamic interactions between data, organisational systems, and individual health care professionals.

The author would like to thank all of the interviewees, both within the health care provider organisations and elsewhere, who gave so generously of their time and expertise. During the development of this research Huw Davies was a Harness Fellow in Health Care Policy at the University of California, San Francisco (UCSF). Thus, this work was supported by The Commonwealth Fund, a New York City based private independent foundation. However, the views presented here are those of the author and not necessarily those of The Commonwealth Fund, its directors, officers or staff. Huw Davies is sincerely grateful to The Fund and the Institute for Health Policy Studies (UCSF) for the opportunities afforded to him by the Harness Fellowship. In addition, Alison Powell assisted with some of the transcript analysis, for which the author is duly grateful.