Health Quality & Safety Commission New Zealand
Kupu Taurangi Hauora o Aotearoa

EVIDENCE REVIEW AND APPENDICES

Position paper on the transparency of information related to health care interventions

Acknowledgements

The Health Quality & Safety Commission would like to acknowledge the Ministry of Health for their support and particularly for leading and organising the consultation forums surrounding this paper, and the Ministry of Health library for their assistance.

Published in March 2016 by the Health Quality & Safety Commission,
PO Box 25496, Wellington 6146, New Zealand.

ISBN 978-0-908345-24-3

This document is available on the Health Quality & Safety Commission website: www.hqsc.govt.nz

HEALTH QUALITY & SAFETY
COMMISSION NEW ZEALAND
Kupu Taurangi Hauora o Aotearoa

open
FOR BETTER CARE
Hāpai ake te toiora

Contents

Evidence review

TRANSPARENCY OF INFORMATION RELATED TO MEDICAL/SERVICE INTERVENTIONS

> I believe the reality is with little or no choice in the matter, the consumer must trust 'the system' to monitor all aspects of a surgeon's competency and that the information that is gained be used by all district health boards to assist with best practice models throughout the country so having a heart attack in Whangarei will give me the same opportunities to survive as having one in Dunedin.
>
> *(Submission from Kevin Salmon, Chair, Northland District Health Board Consumer Council – see Appendix 3)*

This paper was written in response to an Official Information Act request for data on the performance of individual surgeons. However, the principles apply more widely, and across all medical specialties. Lessons from the United States (US), the United Kingdom (UK) and elsewhere are useful to considerations of the approach New Zealand should take.

Public reporting has its origins in Florence Nightingale's work on hospital mortality rates in 1863.[12] Inappropriate variation in the management of patients is a major problem in health care.[96-105] The New York State Cardiac Surgery Reporting System (NYS CSRS)[10, 62] was initiated in 1988 as a response to fivefold variation in hospital mortality data. It is the first and longest-running programme of public reporting of surgical data, and the most studied. Such public disclosure of health care performance data could contribute to reducing variation, driving quality improvement and improving consumer choice, transparency and accountability.[43, 106] Indeed, Shekelle et al (2013) include surgical outcome 'report cards' in the top 10 patient safety strategies for all organisations.[107] However, the primary question is not whether data on the outcomes of surgical and other procedures should be reported, but how this reporting should be done to ensure it is credible, fair and effective in achieving these objectives.[108]

Why transparency?

Arguments for transparency of information about quality of health care depend upon two contentions. First, that transparency is of itself a good thing; second, that it has good results (ie, it in some way leads to better health care).

The first contention is ultimately a moral judgement and as such is untestable empirically. It has wide support, including the support of the Health Quality & Safety Commission (the Commission). However, it is possible that the practical effects of publication (or at least the effects of publishing in a particular way) may cause more harm than good, on balance.

What does publication do?

Publication may change behaviours by patients or providers (whether organisations or individuals). These changed behaviours may have positive and negative effects.

Patients may:

- use the information to change provider, increasing the use of the 'better' providers, and stimulating lower performers to improve or leave the market
- use the information to increase their voice in the system
- gain trust in the system, leading to reduced anxiety and increased compliance with treatment regimens, or conversely, lose trust in the system, leading to increased anxiety and reduced compliance with treatment regimens

- respond inconsistently. More affluent, educated, mobile, and socially powerful patients may use the information to greater advantage than others, which could increase inequity in access and outcomes.

Providers (whether individual or institutional) may:

- respond by improving quality of service because of intrinsic motivations, professional pride, organisational reputation, market mechanisms or financial incentives
- cease to provide a service. This may be desirable, but could reduce access to needed services that are in fact of adequate standard even if not the best in New Zealand. This is especially relevant in more isolated, rural areas where access to many services is already more difficult
- respond perversely by reorganising services to reduce exposure to riskier patients to give the perception of enhanced performance
- respond perversely by focusing on data recording practice rather than improving services
- respond by increased assertion of the importance of individual responsibility to the detriment of a focus on teamwork, with the perverse potential for increased variation in practice and poorer outcomes overall (this theme is discussed further below).

The following section reviews the evidence.

CONSUMER CHOICE AND AUTONOMY

Introduction: Vote with your feet and informed consent

One of the standard arguments for the utility of publicly reporting performance data is that patients may use this information to change provider, increasing the use of the 'better' providers and stimulating lower performers to improve or leave the market.[36, 109, 110] However, publicly reported information may not drive a free-market mechanism where patients 'vote with their feet', even in a US market-driven situation.[34, 43] Furthermore, patients' choices of doctor are influenced by other factors than outcome data, especially for low-risk procedures but also for major surgery. These include the advice of friends and family, access and proximity, and their relationships with particular doctors.[43, 111, 112]

There is, further, an ethical argument for providing patients with comparative individual surgeon information, on the grounds of patient autonomy and the requirements for informed consent to any medical intervention.[13, 27, 113] This argument is not deactivated by some patients' lack of or reduced ability to act on this information (eg, through limited ability to choose their surgeon, as in New Zealand's public health service). There is almost always the possibility of choosing to decline treatment, and autonomy requires that patients understand what they are consenting to.[114]

However, opinion is divided on whether data on surgeon-level performance is a requirement for informed consent. There is some evidence that surgeons who undertake higher numbers of procedures get better results, although the relationship may be correlational rather than causal.[16-18] In Wisconsin in 1996 (in Johnson vs Kokemoor) the state Supreme Court found that a surgeon's experience and risk statistics are relevant to informed consent.[23, 27] However, the contrary was found in Texas in 1996 (in Avila vs Flangas, where it was held that not disclosing information on volume was not a breach of informed consent as inexperience is not a risk 'inherent to the procedure')[24, 26] and in Pennsylvania in 2001 (in Duttry vs Patterson, where it was held that a surgeon's experience was outside the scope of information required for informed consent).[25, 27]

As of 2012, evolving US case law suggests information on surgeons' volume of practice is only relevant in cases of 'substantial inexperience or disadvantage'. The test is whether a 'reasonable person' in the particular circumstances in question would consider the information material to informed consent. In the UK the courts have similarly moved from a standard determined by the profession to one determined by 'the prudent patient' in the particular circumstances in question. In other words, the matter needs to be decided on a case-by-case basis rather than in a universal or general way. For example, one prudent patient may simply want to know that his or her team is safe, effective and compassionate, whereas another might want to know the comparative data on outcomes.[28, 29] Volume and risk-adjusted mortality rates for selected specialties are available on the National Health Service (NHS) Choices website.[115]

In New Zealand, informed consent has its ethical basis in the principle of respect for autonomy. The Code of Health and Disability Services Consumers' Rights 1996 (the 'Code of Rights')[30] is the key source of law on informed consent supplemented by other legislation and case law. The applicable rights are 5, 6 and 7 as follows:

- right 5 addresses a patient's right to effective communication
- right 6 relates to a patient's right to be fully informed; it does not require that the patient requests the information
- right 7 addresses that the treatment can only be provided to a patient if that patient makes an informed choice and gives informed consent, unless an exception applies right 7 also sets out a competent patient's right to refuse or accept any treatment.

Informed consent is an ongoing process involving continuing and appropriate sharing of information throughout a patient's treatment, accompanied by the provision of suitable advice. It is not primarily about compliance or filling out forms. Patients must be given enough information about the proposed treatment and alternatives (including the alternative of doing nothing), to make an informed decision. The information should cover expected risks, side effects, benefits and (if applicable) costs. Honest and accurate answers should be given to any questions including questions about the identity and qualifications of the provider of treatment.

The legal position is clear in New Zealand: the test is that of the prudent patient in the circumstances of that particular patient. Does this imply that volumes and rates of complications are required under New Zealand law for informed consent? Possibly, in some circumstances. However, the material issue may be more general – for example, it may actually be more useful for a patient to know that the team as a whole is performing sufficient numbers of cases and achieving acceptable outcomes, and that mechanisms are in place to ensure that all members of the team are maintaining acceptable standards.

This may be best achieved through a general, national process rather than through expecting patients to check out each service at times of need (and vulnerability). Trying to make sense of the actual data about each individual or even each unit may not be easy for all patients at all times, although it might be for some, at some times.

At a time of need, given a trustworthy overall system, many patients may still like to know the name of the person who will actually provide the surgery or anaesthesia and whether he or she is a specialist or a trainee, and in the latter case some information on how this will be supervised (for example, at night, much work is done by trainees under distant supervision).

Could disclosure impact on training?[16] UK evidence has shown publication of surgeon-specific data has coincided with a decrease in the proportion and variety of cases performed by trainees, alongside increasing complexity of casemix. This could reflect decisions by patients but it could also reflect decisions by specialists keen to ensure the cases attached to their names do well.[116] However, the Commission's view is that gaining experience cannot be more important than informed consent, and if training is properly supervised there should be no material increase in risk.

The Royal New Zealand College of General Practitioners (RNZCGP) supported the use and publication of outcome data, but advised careful consideration.[117] They noted an important risk for generalists. General practitioners could decide to focus on subspecialties to improve their reported outcomes. Experienced generalists are important for diagnosis of undifferentiated patients, so this could reduce safety in one way while increasing it in another.[117] Access to essential non-specialised services might also be reduced.

The RNZCGP further notes a potential 'subtle and pervasive impact' of outcome reporting on informed consent and patient choice; that of treatment variation: 'If a patient has a choice of interventions, and one is higher risk than the other, the implications for… this will clearly impact on the patient's ability to make an informed choice.'[117]

Further, it is relevant to ask that if a surgeon's performance is determined by the institution and team within which they work, and that performance alters when they change team or institution, as Huckman and Pisano (2006) found, to which 'version' of the surgeon are patients giving their consent?[31]

In any case, it has been argued in the Australian context and elsewhere that the inaccuracy of the surgeon-specific performance metrics obviates the ethical obligation to disclose them as part of the process of informed consent – including in terms of the substantial time lag for validated data to emerge.[27, 32]

Individual surgeon 'report cards' require the allocation of substantial resources to ensure adequate risk adjustment and accuracy, and 'to justify such a project, we need to show that the benefits outweigh the costs.'[13] Assuming valid data are produced and published, how then are they used by consumers and what benefits do we see?

The evidence

A 2015 Cochrane review of the effect of provision of performance data for people considering elective surgery found no qualifying studies.[33] A 2011 Cochrane review of the impact of public reporting on consumer choice found four qualifying studies, which, despite containing more than 35,000 consumers and 1560 hospitals in North America and Canada, provided no consistent evidence that public release of performance data changed consumer behaviour.[34]

Do consumers use the data?

The evidence is, in general, somewhat mixed, but largely suggests they don't.

In a review of the literature between 1986 and 1999, Marshall et al (2000) suggest that public reporting had limited impact on consumer decision-making. Less than 25 percent of consumers stated that performance data had any significant impact on their choice of surgeon, and consumers continued to use hospitals with high mortality rates.[42] Reasons cited include difficulty in understanding the information, disinterest, lack of trust in the data (especially those publishing it), problems with timely access, and lack of choice.[42] Four years later they found that although consumers are in favour of public reporting, in the US they tend not to search for the information, sometimes fail to understand it, mistrust the quality of the reports, and make little use of the reports in actual decisions.[41]

Awareness is a major issue. A 1998 study on the use of a statewide consumer guide providing risk-adjusted mortality ratings of hospitals that provide cardiac surgery found that only 12 percent of those surveyed were aware of this resource. Fewer than 1 percent knew the correct rating of their surgeon or hospital and reported that it had a moderate or major impact on their decision.[36] In 2015, however, a US study showed consumers did use online quality reports, although at lower levels than the UK's NHS Choices website. They found that US consumers predominantly searched for individual hospitals rather than individual practitioners or by clinical condition and were most interested in patient narrative comments (31.5 percent, 139/442).[118]

Emmert et al (2013) showed that physician-specific reporting websites in Germany played a role in consumer decisions – 65 percent of consumers who had used such a site (249/381) consulted with a particular physician based on the website ratings.[119] Grabner-Krauter and Waiguny (2015) found fact-based reviews were seen as more important than emotion-based reviews when there were a low number of reviews.[120]

Despite the growth in public reporting activities over the past 27 years, there is limited evidence of their use by consumers in ways that have significantly affected health care delivery.[47] This may reflect flaws in the content, design and implementation of existing public reports rather than inherent limitations of reporting per se.[46] Support for public reporting continues, in part, because of the face value of transparency, but substantial work is still needed for public reports to achieve their potential for engaging and informing consumers.[45]

Providers' views

Mannion and Davies (2002) interviewed 18 US health experts who said that public reporting had raised the visibility and consumer awareness of variation in quality and performance.[121] They felt that usage remained poor due to lack of input from potential users, and the public had low confidence in such information. Also, choosing alternative providers often required unfeasible travel.[121]

What is known about presentation of information for the public?

Publicly reported quality information, including surgeon performance data, is generally presented in ways that consumers find difficult to understand. The RNZCGP notes that

> Publicly available data must be put into the appropriate context … this is also a requirement on healthcare providers under Right 5(1) of the Health and Disability Consumers' Code of Rights: 'Every consumer has the right to effective communication in a form, language, and manner that enables the consumer to understand the information provided.' Our interpretation of this requirement is that it would be inappropriate for a health service to provide raw outcome data if that data is likely to be misinterpreted by a health consumer. That does not mean that the health service should not provide the data, but instead that it must ensure that it is provided in a way that can be understood.[117]

The RNZCGP uses the argument that more-experienced practitioners are dealing with cases with greater risk profiles, and thus have ostensibly poorer outcomes. Whether or not under the Code of Rights practitioner-specific outcomes were intended to be construed as such 'communication', the issue of presentation of such data and its comprehensibility by consumers is a vexed one, as we shall see.

Donelan et al (2011) investigated the format of presentation of surgeon performance data for coronary bypass procedures and how this might affect consumer comprehension and thus choice of cardiothoracic surgeon. Across four different presentations of the same data, 'only 6.4% [of consumers] identified the surgeon with the lowest risk mortality though they felt it important' to find this information.[38] This was echoed in a 2009 systematic review of 14 studies, which found that format of presentation was important to improve consumer knowledge, and thus choice.[39]

Robinowitz and Dudley (2006) felt that two potential explanations for the low level of impact are that, in most early reports, the large majority of providers were labelled 'average' and consumers may have had difficulty understanding the statistical assessments.[40] Reporting of outcomes data needed to match the health literacy of the population, or else the data might be too complex (or too vague) for prospective patients to make an informed decision.[122] Sinaiko et al (2012) interviewed key stakeholders including consumers about public reporting, style of publication and where reporting should go in the future.[6]

All stakeholders did not doubt the value but overall felt the reporting had become disconnected from consumer decisions and that reporting needed to change to realise the potential to inform consumer choice.[6] This included the adequacy of quality measures and the need for functional and health outcomes rather than mortality alone; composite measures tailored to segments of the population by demographic and health status; and tailored reports for individuals, especially for those who want to drill down creating more personalised customisable reports to meet the needs of individuals.[6]

The reports need to be formatted, presented, written and delivered in different ways, and funding needs to support data collection in this way.[6] In designing public reports, Bridgewater et al (2013) emphasise that patient representatives are essential in the discussion and that reported outcome measures must be important to patients but also clinically relevant and measurable.[4]

Accuracy and validity of data must also be considered. Davis et al (2007) report that consumers had low tolerance for inaccuracy in rating individual physicians.[123] The agency reporting the data is an important factor determining its acceptability, especially if a government or political agenda is perceived.[41]

Minami et al (2015) confirm that current metrics for surgical outcomes aren't patient-facing, aren't individualised, don't speak to different kinds of people with different conditions and different priorities (other than mortality), and are hard for consumers with low literacy or numeracy to understand.[7] The authors posit 'a patient-specific tool that takes into account individual risk factors and demographic information so that patients can compare their own peri-operative risks at different hospitals that accept their insurance', potentially as an 'individual composite score'.[7]

Friedberg et al (2012) present a five-point methodological checklist to guide those who want to improve their performance reporting methods with the goal of helping report makers minimise the frequency and severity of provider performance misclassification and avoid adverse unintended consequences of reporting.[124] The checklist directs those who produce reports publicly to address and publicly explain each checklist item, thus increasing transparency and encouraging more rigorous methods to improve the chances that reports will lead to better, more efficient care.[124]

The authors believe public reports of provider performance on measures of the quality, costs, and outcomes of health care can spur improvement and help patients find the best providers, but that the likelihood that these benefits will materialise depends on the methods underlying each performance report.[124]

Though sparse, there is, however, evidence of effects when thoughtful presentation of information is involved.

Hibbard et al (2005), in a study of the Wisconsin QualityCounts programme, found that consumers were able to more accurately discern between higher and lower quality hospitals based on these reports, and that these effects persisted over two years.[48] Considerable thought had been put into the presentation of the QualityCounts reports. They found the influence of this finding was more likely to be on provider reputation, however, rather than market share[48] – the potential provider effects were more likely to be on US-specific arrangements, like fundraising and donations. Consumers were more likely to be concerned about negative performance findings than 'provider shopping' – we discuss the effects on reputation as a driver for quality improvement below.

Furthermore, Godlee (2012) describes how, when the Barnsley Primary Care Trust in the UK assessed all its general practices against agreed best practice for 13 common conditions and published good performers with a simple and easily comprehensible green tick logo, between 5000 and 7000 patients changed practice as a result.[15]

Effects on equity – ethnic and socioeconomic

Public reporting of specific quality information such as individual surgeon 'report cards' appears to have potentially negative effects for equity.

Fung (2008) et al in a systematic review showed that public reporting led to increased disparity amongst racial groups due to mechanisms of inequitable access to and use of such reports, which did reduce over time, and that higher socioeconomic level neighbourhoods were more likely to be treated by surgeons with low risk-adjusted mortality rates, whereas the reverse was true in lower socioeconomic neighbourhoods.[125] Shahian et al (2011) found the effect of publicly reported outcomes on referrals is modest, transient and generally limited to more affluent and educated areas.[12]

Christianson et al (2014), in a review in the US Medicare context, found low awareness of publicly reported quality information among older people with chronic illness, thus limiting their decision-making based on this information and potentially increasing disparity in regard to vulnerable groups.[49] Other papers have explored this potential marginalisation effect, especially in terms of ethnic inequalities becoming exacerbated after public reporting.[50, 51, 54] In a New York study, Werner et al (2005) found different usage rates for coronary artery bypass graft (CABG) surgery in black, Hispanic and white patients, and hypothesised that surgeons in a publicly reported outcomes environment were avoiding CABG surgery for black and Hispanic patients based on those patients' higher risk for poorer outcomes.[54] Some have argued that including such information as part of the process of seeking informed consent would circumvent the issues of awareness and timely access.[27] It has been shown that once minorities are aware of reports, they are more likely to use the information.[53]

In their submission on the draft of this paper, the Accident Compensation Corporation mentions the potential negative effects on equity and rural service provision:

> The greatest [moral] hazard [of publication] relates to the likelihood that any benefits would most improve health outcomes for those who are already advantaged.... [D]emographic factors such as poor outcomes from already disadvantaged patient populations may well reduce the participation of providers... in some institutions, especially rural ones, or to select low risk patients, less likely to benefit.[126]

Casalino et al (2007) write, 'Unless carefully designed, these [public reporting] programs may have the unintended consequence of increasing racial and ethnic disparities' as well as socioeconomic disparities.[127]

What drives consumer choice?

Schlesinger et al (2014) describe consumer choice as a 'black box' – consumer decision-making with regard to physician choice could not be explained.[44]

Several studies have shown that consumer choice is not driven by information on quality.[35] In reviewing the impact of the publicly published Scottish clinical indicator programme, Mannion and Goddard (2003) interviewed consumer councils, who felt that consumers rarely sought out the data.[37] Most consumers were felt to make a choice based on family and friends' experience, on their own past experience, and the advice of their general practitioner (GP). Although 25 percent of GPs consulted the indicators, GPs based their decisions less on the indicators and more on waiting times and local audit reports. GPs rarely discussed the indicator data with patients.

US evidence confirms that consumers rely more on anecdotal evidence from family and friends than empirical evidence.[109] Morsi et al (2012) looked at the use of websites aimed at health care in the US and showed that in 2008 only 14 percent of Americans had used such information in the past year, and that after seeking input from family and friends they generally relied on their primary care physician to assist them in making decisions about where to have elective surgery.[11] Studies also show that cardiologists don't use publicly reported quality data in their referrals to cardiac surgeons, and again there is limited discussion with patients.[128, 129]

Shahian et al (2011) write that, in terms of informed consumer choice, report cards were ethically desirable in terms of enhancing the decision-making autonomy of patients but have not been used to the extent envisaged by health policy experts. Most patients were unaware of their existence, did not know how to access them, lacked trust in them or did not know how to interpret the ratings.[12] The reputations of hospitals and surgeons were still the most important considerations for patients selecting their surgeon, and many relied heavily on the advice of their primary care doctor.[12] Highly publicised adverse events or poor score card ratings have a greater influence then identifying top performers.[12]

New Zealand consumer feedback to the draft paper raised that choice is not a reality for the vast majority of the New Zealand population, and the fact there isn't a choice should be all the more reason for complete transparency to drive improvement across the sector. Further, if the concept of patient choice is illusory then data release will be of limited utility unless it constitutes a tipping point for declining surgery.

Performance estimates must be credible, risk-adjusted, reliable, discriminating amongst providers, and presented in a manner that can be accurately interpreted.[12] Shahian et al suggest that whilst professional organisations and societies have a special responsibility and right of self-regulation (see 'Accountability' later), they are obligated to put the interests of patients first.

The authors suggest professional organisations and societies could make a unique contribution by providing national risk-adjusted benchmarks against which to gauge provider results and are best qualified to provide credible and accurate information best addressed through complete transparency of the methodology and audit processes they use.[12] The authors state there is strong and consistent objective evidence supporting the positive impact of performance measurement, but the incremental effect of publicly reporting is less certain, and it is unclear if public reporting is the only or best way to achieve quality improvement (see 'Quality improvement and public reporting' later).[12]

What measures should be reported?

How do we make data salient for patients? Rothberg et al (2009) suggest we should prioritise common elective procedures where choice is possible, include quality of life and outcome measures that are risk-adjusted, and include measures of patient experience such as satisfaction.[130] Careful attention must be given to visual representations that convey relative difference, and those due to chance must be identified and communicated. The use of non-validated administrative data should not be undertaken, and if the data is not available there is a cost in setting this up. Whilst the paper suggests that reporting should be patient-centred, it recognises this as challenging to implement. Process measures and mortality rates remain abstract concepts; what patients value most is knowing the experiences of others, whether pain will be controlled, whether they will be listened to and whether the nurse will come when they call.

There is a robust argument to be made that publication of 'hard' clinical outcomes are by themselves insufficient to drive improvement of quality and ensure care is delivering outcomes that matter to patients. Patient reported outcome measures (PROMs) – defined as an individual's assessment of their health or wellbeing that comes directly from the individual without interpretation by a clinician or anyone else – are an important adjunct to clinical outcomes. Most health care aims to reduce symptoms, minimise disability and improve quality of life – aspects that only patients can assess.[131] The promise of PROMs has been noted for at least 25 years, but the impact to date on clinical practice has been limited.[9] How best to use this potentially powerful tool is a question of wider scope than we have here, so we limit our comments to noting this as a useful area for further development.

In a systematic review of 27 studies, Chen et al (2013) found PROMs that were not publicly reported showed evidence for improved patient–provider communication and patient satisfaction, but not for improved outcomes or effectiveness of quality improvement programmes.[8] In the NHS in England, patients are encouraged to provide feedback to hospitals on their quality-reporting website, NHS Choices.[115]

Lagu et al (2013) analysed feedback in all NHS hospitals with more than 10 reviews posted (264 hospitals and 200 reviews with 2640 patients surveyed).[32] In 124 (62 percent) of the 200 reviews, patients commented on technical aspects of hospital care, including quality of care, injuries, errors, and incorrect medical record or discharge documentation. Perceived medical errors were described in 51 hospital reviews (26 percent). Comments about the hospital facility appeared in half (52 percent) of reviews, describing hospital cleanliness, food, parking and amenities.

Reviews appeared to have similar domains to those covered in existing patient satisfaction surveys but also included detailed feedback that would be unlikely to be revealed by such surveys. Lagu et al concluded that online narrative reviews can therefore provide useful and complementary information to patients and hospitals, particularly when combined with systematically collected patient experience data.[32]

Many district health boards (DHBs) have been conducting their own patient satisfaction surveys for many years, and since August 2014 the Commission has conducted a systematic nationwide patient experience survey in New Zealand hospitals.[132]

New Zealand consumer needs – do New Zealand patients want information about a surgeon's performance?

At a recent workshop held by the Commission and the Ministry of Health in June 2015, 24 consumer network representatives echoed many of the points raised in the research noted above.[5] Consumers wanted useful, relevant, quality information that was easy to understand, accurate and valid, and included both private and public systems.

They wanted information from a consumer perspective centred on the patient journey, such as wait times and cancellations, with data on two to three key aspects of a procedure, published on each DHB website. They wanted details of the process and likelihood of outcomes including quality of life, and the risks and benefits for themselves as individuals. They wanted opportunities for stories to come through a mix of data and personal accounts – as well as patient experience surveys and the ability to access 'expert patients' who had had first-hand experience. Consumers stated that 'loads' of numbers were not helpful, especially given the level of choice that is available. Information needed to be valuable – not '30-second soundbites'. They wanted many sources and the information to be presented in many ways, such as DHB websites and social media, and published in consumer-friendly formats.

New Zealand consumers wanted trust and confidence in the system and to know that the professionals were competent and met or were above standard, and that this competence was publicly demonstrated (see 'Accountability' later). They wanted to know if a surgeon is the best qualified for the surgery to be undertaken, whether they have had adverse events or investigations, and had done the volumes, though they realised and acknowledged that the more granular the information the more distorted it became, and that care was a 'team game'. In light of this they wanted to know that the organisation had professional teams, there was effective team briefing/debriefing, and that surgeons were connected to the wider multidisciplinary team and other professionals, made plans together with them, and referred appropriately to others.

Consumers also wanted to know that the system was reliable and the organisation had the right culture, that there was a visible process to improve quality, issues were being addressed openly and transparently. They wanted to know that the institution met standards and had organisational safety nets, and an active system for learning and improving – sharing the positives as well as negatives – and acted when things weren't right. They wanted to know there were systems to keep them safe and a focus on quality improvement. Consumers wanted to know that this information and these processes were seen and linked to public board reporting of quality improvement and safety data. Consumers wanted to be active and engaged at all levels of the development/design of such a system. Clinicians at previous workshops said measures need to be co-designed by the centre with clinicians and consumers.

Sector feedback has noted that 'consumer representatives do not view or use health care data in the same way as the people and populations who use health care, and evaluating the likely benefits [of publication] (in terms of accessibility and health benefits) would require a wider public consultation process.'[126] Nevertheless, we feel feedback from the consumer workshop is in line with the evidence presented above.

Consumer feedback stated that it should not be assumed patients will not value such data nor should it be assumed that only a small proportion of patients would use such data. An appropriate cost-benefit analysis is an important prerequisite. Surgical outcome data around quality, not just mortality, is useful to patients and clinicians, bearing in mind the limitations of choice.

DATA LIMITATIONS AND IMPACT ON PATIENT AUTONOMY AND CHOICE

Ideals and realities

The delivery of health care is a complex enterprise. This makes creating a set of indicators that are both useful and readily understood by the general public very difficult. If outcomes are published to enhance patient autonomy and give choice, the data must be accurate, valid, and measure the outcome stated. A focus on mortality data alone does not measure physician quality, nor is it what patients themselves want.[1, 5, 112, 133]

Murphy (2012) looked at good practice in measuring and reporting health outcomes and discusses the complexity in measuring quality and outcome; the difficulty in deciding the criteria for quality; the temporal relationship for any measure; and for outcome data to be risk-adjusted.[134] The ideal measure needs to be validated, specific, sensitive, reliable, responsive, timely and easy to measure.

The challenge for the profession, consumers and the specialty groups is to agree which quality indicators should be reported and whether the data are collected at a clinician-, care pathway-, team/department- or organisation-specific level. This then needs to be risk-adjusted and entered into national data sets. Measuring this requires agreed standards of data collection across New Zealand, the development of agreed risk adjustment models and data definitions, and bringing all the data into a central collection system, which needs to happen as part of business as usual.

Data on a national scale are expensive: provision and maintenance of the cardiothoracic surgical data system in the UK costs £1.5 million per year. The cardiac registry in New Zealand costs $1 million per year (K Evison, Ministry of Health, personal communication, 3 July 2015). Auras (2012) found that German hospitals felt the cost-benefit ratio of mandatory public quality reporting was too high. The burden needs to be considered.[135]

The Society for Cardiothoracic Surgery in Great Britain and Ireland (SCTS) national database, which collects data on all adults undergoing cardiac surgery, began in 1994 with 12 hospitals and now takes data from all NHS cardiac surgery units.[136] The main aim of the database was to develop reliable, UK-orientated risk stratification models. An early study designed to assess the quality and completeness of the SCTS database in 2003 'revealed it to be both incomplete and unreliable in its ability to yield accurate, risk-adjusted outcomes data'.[13, 137] The process of development required national cooperation and research and innovation over a considerable period of time.

The SCTS maintains that 'routine collection, benchmarking and feedback of clinical outcomes… reduces the costs of health care delivery'.[20] Indeed, using international cardiac surgical benchmarking they show that patients in those countries that actively benchmarked and fed back clinical audit data have a shorter length of stay. At only one day shorter length of stay, 20,000 CABG operations in England per annum equate to savings of £5 million a year – more than enough to cover the initiative.[20, 108] It is a lengthy and complicated process, but in the long run may be cheaper. Present data on surgical outcomes in New Zealand are administrative, not risk-adjusted, and not collected for the purpose of reporting surgical outcomes.

Power limitations and false complacency – 'absence of evidence is not evidence of absence'

Walker et al (2013) have explored the potential dangers in regard to statistical power of data in the specific context of surgeon-specific outcome reporting in English NHS hospitals.[14] The danger is of lower numbers of procedures masking poor performance leading to false complacency – that is, that surgeon-specific mortality is an inadequate and potentially misleading measure of surgeon performance overall.[14]

The authors asked: What number of procedures is necessary to reliably detect poor performance by a surgeon? Further, how many surgeons would actually perform this number of procedures over a period of one to five years, and what is the probability that a surgeon identified as a statistical outlier truly has poor performance?

The authors used the UK national postoperative mortality rate for select periods and procedures, and defined poor performance as double that rate – 'in practice… a fairly large difference in performance'.[14] The number of procedures surgeons must undertake annually to reach acceptable levels of statistical power to detect eight in ten true poor performers exceeded the median number of procedures typically undertaken by nearly threefold for cardiac surgery, more than threefold for hip fracture surgery, 13-fold for oesophagectomy or gastrectomy, and 20-fold for bowel cancer resection. In absolute numbers, 132 to 179 bowel cancer resection procedures would be required annually for 70 to 80 percent power. The mean annual number of procedures performed in the UK is around 9.[14]

An option is to pool numbers over a longer reporting period. Three-quarters of surgeons in the UK do sufficient numbers of cardiac surgical procedures to achieve 60 percent statistical power over a three-year reporting period, and the same for hip fracture surgery. There are thus consequences in terms of timeliness and relevance, which could mask recent deterioration. Furthermore, even over a five-year reporting period only a third of surgeons performing bowel cancer resections and gastrectomies or oesophagectomies will do enough procedures to reach this power.[14] Such data are therefore underpowered for accountability and research purposes; too late for quality improvement purposes; are little used by health care consumers; and, if used, risk undermining the public interest by generating false complacency.

Walker et al (2013) suggest that in the specialties in which most surgeons do not achieve 60 percent power the unit of reporting should be the team, hospital or trust.[14] As well as the benefit to statistical power, Chou et al (2015) state it is 'important that team-level data are published as well to reflect the complex interplay of the multi-disciplinary team'.[138]

Walker et al (2013) suggest results should be presented as funnel plots, and no interpretation of poor performance should be given until further investigation is made. Equally, no evidence of poor performance should be interpreted as evidence of acceptable performance.[14]

It should be noted that funnel plots themselves may be misleading – the results of an individual (or unit) with a low number of cases may be worse than those of one with a high number of cases, yet the former individual may fall within the confidence limits considered acceptable on the funnel plot while the latter individual falls outside them, and is flagged as a poor performer. A proper interpretation would be that there is simply not enough data to make a reliable assessment of the former individual or unit, but this explanation is demonstrably missing from most publicly reported data that are presented in the form of funnel plots.

Chou et al (2015) hold that despite the controversy surrounding the topic of publishing surgical outcomes, the advantages of reporting outcomes outweighs the disadvantages. The authors stress that cooperation between the central health agencies, royal colleges, professional societies and patient representatives is needed to collect national data regarding performance quality.[138]

However, the difficulties and challenges (and costs) of data notwithstanding, as Sir Bruce Keogh (2008) put it in his presidential address to the SCTS, '[Technical] shortcomings are not important in the grand scheme of public disclosure. The point is that the genie is now out of the bottle, there is no going back.'[1]

ACCOUNTABILITY

Many clinicians are worried that publication of raw data of mortality and complications will lead to 'public crucifixion' of the individual surgeon identified with a high mortality rate, who is either misrepresented by the data due to their high-risk casemix, or who is already under review.

As Keogh (2008) says: 'We were concerned by the potential, unnecessary humiliation of surgeons whose mortality may be high. At worst this would result in the public crucifixion of a competent surgeon; at best it would further humiliate a surgeon in difficulty, who with the current systems in place will already have been identified, and hopefully helped.'[1]

But would that surgeon in difficulty already have been identified and helped in New Zealand? What systems are in place? Don't prospective patients of that surgeon have a right to know?

Current assessment of professional competence in New Zealand and developments underway

Professional competence of medical practitioners is complex and has been defined as 'the habitual and judicious use of communication, knowledge, technical skills, clinical reasoning, emotions, values and reflection in daily practice for the benefit of the individual and community being served'.[139]

Competence is what a health care professional has been trained to do, whereas performance is what they do in daily practice, and is thus influenced by both individual and system factors. Addressing issues of competence is a continuous process throughout a medical career, from postgraduate training to specialist practice. Doctors need to demonstrate both clinical competence (technical skills and knowledge) and behavioural competence (interpersonal and affective skills, such as the ability to communicate effectively, use judgement and empathy, and manage relationships).[140] The Health Practitioners Competence Assurance Act also requires doctors in New Zealand to be 'culturally competent'.[141]

A number of separate but overlapping processes are currently used to demonstrate a doctors' ongoing competence. They vary across organisations and across speciality groups but include individual credentialing on taking up work as well as individual annual credentialing, service credentialing, performance appraisals by employers, College Continuing Professional Development processes that support recertification by the Medical Council of New Zealand, regular practice reviews (recommended by the Medical Council of New Zealand and applied by some colleges/associations (Royal Australian and New Zealand College of Obstetricians and Gynaecologists, New Zealand Orthopaedic Association and RNZCGP), peer review, multisource feedback, and other external quality assurance programmes such as those used in pathology and radiology.

The Council of Medical Colleges, the Medical Council of New Zealand and the Ministry of Health are working to better align and strengthen these processes and are developing a draft best practice guide and stocktake of tools and resources to develop processes and systems to demonstrate doctors' ongoing competence. If completed, these activities would enable medical specialists to demonstrate they are actively participating in processes to improve their professional practice (Ineson S, Executive Director, Council of Medical Colleges, personal communication, July 2015), thus continually improving the quality of patient care.

Exworthy et al (2010) note that the purposes and motivations for disclosing clinical performance are multiple and may include improving (individual and/or organisational) performance, identifying 'poorly' performing individuals or organisations, aiding user decision-making (as part of 'choice'), promoting transparency of professional activities, and securing accountability for public spending.[142]

We can consider accountability in different ways. Although public reporting is often stated to be useful in terms of accountability to the public, we have seen that the public do not search out these reports, often do not understand the presentation, and make little use of the reports in their decision-making as choice is often not available. Where choice is available, friends, family and GP advice are often factors, alongside the proximity of the care.[43]

Although assessment of practitioner competence does happen at present, the system could be strengthened and made more transparent so the information is readily available and the outcomes visible. Chou (2015) stresses that cooperation between the central health agencies, royal colleges, professional societies and patient representatives was needed to collect national data regarding performance quality in the UK.[138] There is widespread agreement in the literature that the public should be involved in building public reporting systems.[1-4] And we have seen that the New Zealand public themselves wish to be engaged.

Is there correlation between surgeon quality rankings and disciplinary and complaint rates?

Roberts (2006) looked at surgeon rankings of quality as well as disciplinary rates and complaint rates in Wisconsin, and showed that there was no correlation between these.[94] Wisconsin ranked poorer in terms of numbers of license sanctions yet did well by Medicare quality rankings and rates of adverse events.

QUALITY IMPROVEMENT AND PUBLIC REPORTING

The quality argument is that publicly reporting outcomes drives providers to improve quality. But are outcomes such as mortality and morbidity the right information to drive surgeons and teams to improve their performance? What is the mechanism behind improvement? And thus what kinds of data most feed that mechanism?

Quality improvement, accountability or research?

It is increasingly well-accepted that the data required for quality improvement activities differ from that used for accountability or research purposes.[143] In 1997 Solberg et al published their seminal paper distinguishing these three faces of performance measurement, and how they differ in who they are for, their purpose, and what and how they measure.[143]

Measurement for quality improvement is for providers and staff, quality improvement teams, and administrators, Solberg et al suggest. It is or ought to be limited to an individual clinical site and process and the measures ought to be few, easy to collect, and approximate, over a short, current time period, with confounders considered but rarely measured. The sample size is small, the measurers are internal, unblinded and involved in selection of measures. Thus the need for confidentiality is 'very high'.[43, 143] As Seddon (2012) puts it: 'Without this confidentiality, the risk is a loss of trust in the process [of quality improvement measurement].'[43] She quotes O'Leary: 'The problem with measurement is that it can be a loaded gun – dangerous if misused and at least threatening if pointed in the wrong direction.'[43, 144]

The goal differs from that of measurement for research, which is 'new knowledge without regard for its applicability'.[143] Here the scope is universal. There are many measures, which must be complex, precise and valid. Confounders must be identified, measured and controlled. The samples must be large, researchers blinded, and the need for confidentiality is high. This kind of research is thus expensive, elaborate and rarely timely, due to the rigour required to produce 'knowledge of wide generalisability or universal value'.[143] The goal differs also from that of measurement for accountability – where the goal is exposure and comparison, and the data must therefore be valid and reliable, taken from a large sample size over a long period of time, and rigorously risk-adjusted.[43] The goal of measurement for quality improvement is understanding of process, motivation and focus; establishment of baselines; and evaluation of change over time in rapid cycles of learning.[143] There is overlap in that the comparisons made possible by accountability measures that are consistent and risk-adjusted across jurisdictions may stimulate quality improvement work. When the data are trusted, hospitals use them for quality improvement work.[145]

Seddon outlines her twofold concerns about public reporting of surgeon-specific outcomes in a more general sense:

1. Technical problems – There are major issues in risk adjustment – in New Zealand you will be relying on coded data and that is variable around the country (to code all the various comorbidities and other factors that increase a person's surgical risk – some of which are unknown). As soon as you start to compare data then you have to put a huge amount of effort into making that data unimpeachable.
2. Philosophical – If we accept that medicine is essentially a team-based activity, it makes no sense to publish data based on one team member.

   However, I think that there is a certain amount of inevitability about this – I am not against transparency and publishing data, just not sure that this will get us where we are going. Also I think that 'choice' is a red herring. Are the private hospitals going to publish their surgeons' outcomes? (M Seddon, personal communication, 6 July 2015)

(NB: Representation from private hospitals at the 3rd July Consumer Representative Forum on Surgical Data was supportive of the direction towards greater transparency.[5])

The ways measurement may lead to improvement

Berwick et al (2003) described two causal pathways between performance measurement and performance improvement: these are *selection* and *change*.[57]

Selection is based on the economic model of a competitive market where patients behave as consumers. Through selection they reward better providers and force lower quality providers from the market. As we have seen in the 'Consumer choice and autonomy' section earlier, this mechanism does not seem to operate in health care: consumer selection does not lead to improvements in quality of care, even in the US context.[39, 41, 42, 48, 52, 56]

However, there is evidence that public reporting stimulates *change* in behaviours of providers at the organisational level, independent of any economic benefit. Why do they improve? Institutions that reported publicly improved their performance more than those that reported internally, regardless of changes in market share. Improvement wasn't about information providing managerial ammunition to know what to improve. The motivation and stimulus to action was reputational damage, status and professional pride. This change made on the basis of public reporting is known as the 'reputation' pathway, or 'name and shame'.[45, 48, 56-61]

The evidence – public reporting, improvement, gaming and morale

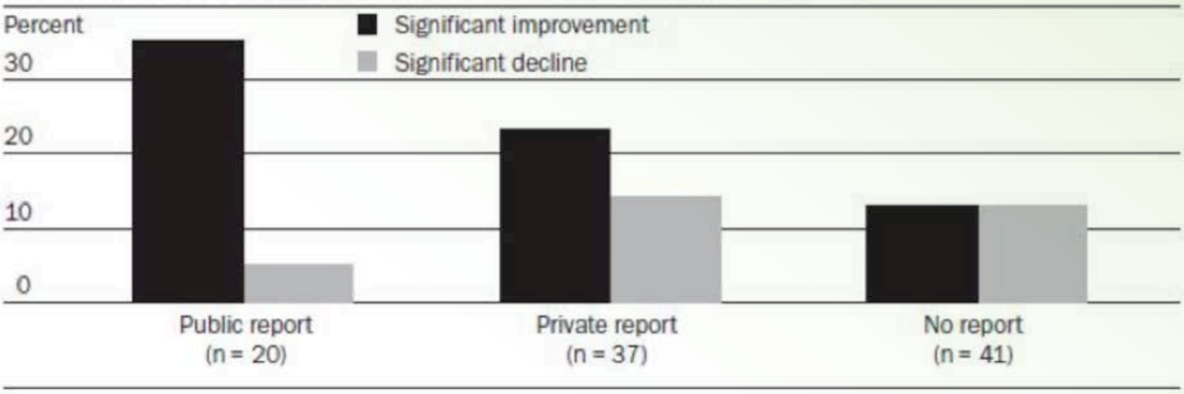
'Reports are not self executive,' Florence Nightingale continually reminded herself.[82] It appears that their proper public presentation is what forces knowledge into action.

Hibbard et al's 2005 study of Wisconsin hospitals is the most widely cited for the positive effects of publication on quality improvement by the reputation pathway. Hibbard et al compared the performance of 24 hospitals in south central Wisconsin, allocated to three groups: those which published the QualityCounts quality measures externally, those who published internally, and those who didn't publish at all. The measures included two summary indices of adverse events occurring within the broad categories of surgery and non-surgery, along with indices summarising three clinical areas: hip/knee surgery, cardiac care and obstetric care. Measures were taken before and after publication. Hospitals were rated as better than expected (fewer deaths/complications), worse than expected, or as expected.

These public reports were very public. They were disseminated widely – copies were distributed by community groups and libraries, inserted in local newspapers, and published on the web. There was press coverage and substantial public interest.

Hibbard et al found hospitals that reported publicly put more quality improvement activities in place, and subsequently showed clear improvements in performance, compared with hospitals that reported internally or not at all (Figure 1).[48]

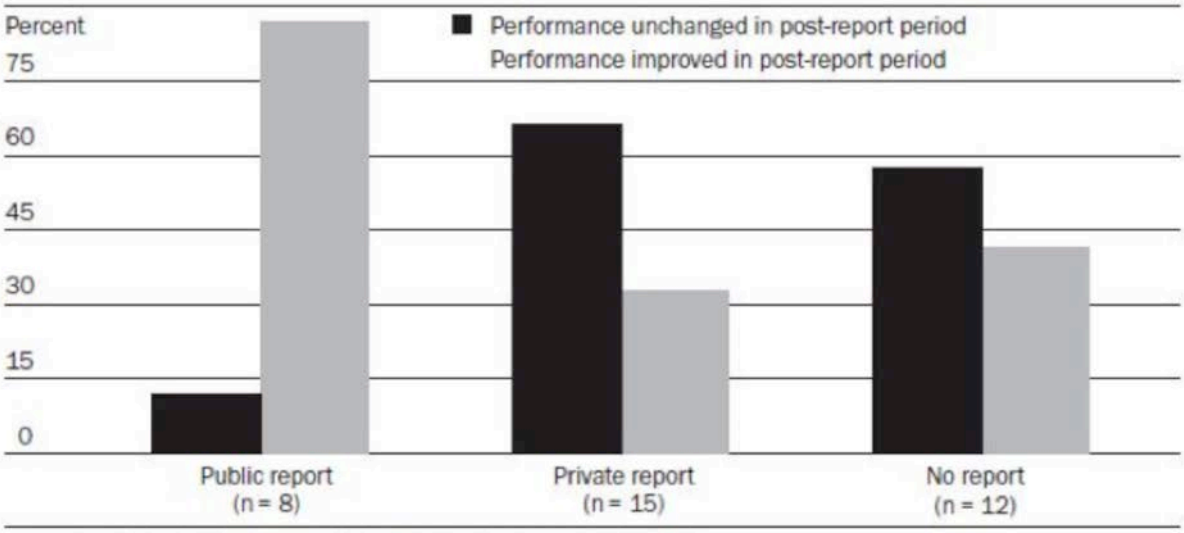
Figure 1. Percentage of 24 Wisconsin hospitals with statistically significant improvements or declines in obstetrics performance in the post-report period (2001–03)



Source: Wisconsin Bureau of Health information, risk-adjusted by Medstat.

Nearly 90 percent of hospitals performing lower than baseline that began publicly reporting their outcomes improved their performance two years after reporting. Of those hospitals performing lower than baseline that reported internally, only 30 percent improved and the performance of more than 60 percent remained the same (similar to the group who didn't report at all) (Figure 2).[48]

Figure 2. Changes in hospital performance in the post-report period (2001–03) among Wisconsin hospitals with worse-than-expected scores at baseline



Source: Wisconsin Bureau of Health information, risk-adjusted by Medstat.

Note: Change in scores at baseline is significant at the 5 percent level.

On this information Hibbard et al hypothesised four characteristics of a performance management system to have an effect. It must be:

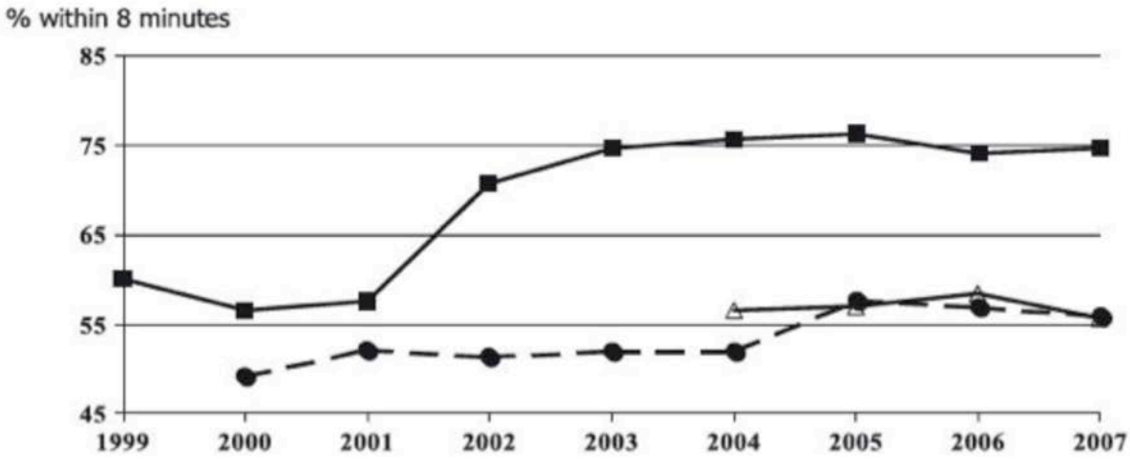
- a ranking system
- published and widely disseminated
- easily understood by the public (so that they can see which providers are performing well and poorly)
- followed up by future reports (that show whether performance has improved or not).[61, 82]

Hibbard et al's paper is widely cited in support of public reporting for quality improvement, but it has its discontents. Shahian et al (2011) point to improvement being limited to obstetrics and areas with already low baseline performance, methodological issues such as low n, minimal risk adjustment and the 'disappearance of statistical significance when appropriate random effects models were used'.[12] Shahian et al's reservations are, however, unreferenced and unpublished, and we contend low baseline performing institutions are exactly those we wish to improve.

Bevan and Hamblin (2009) identify the Wisconsin QualityCounts programme as a strong example of a controlled experiment of the change by reputation mechanism of publication. They then identify a natural experiment that occurred in the UK in the early 2000s when targets were set for, among other measures, ambulance response times. Star ratings – a stark, clearly comprehensible and widely disseminated ranking system based on such targets – were instigated for England in 2001. In Scotland and Wales, similar targets were set with no ranking system, relying on internal pressures alone to meet those targets.[82] In other words, both systems relied on the change pathway, but only one engaged the reputation stimulus by publicly awarding stars to trusts.

The result was that after targets and star ratings were introduced, English ambulance response times improved dramatically to and above the target of 75 percent of category A (urgent) calls being answered by an ambulance within eight minutes. In Wales and Scotland little improvement was seen, and targets not met, even when those targets were reduced in 2004 and 2005 (Figure 3).[82]

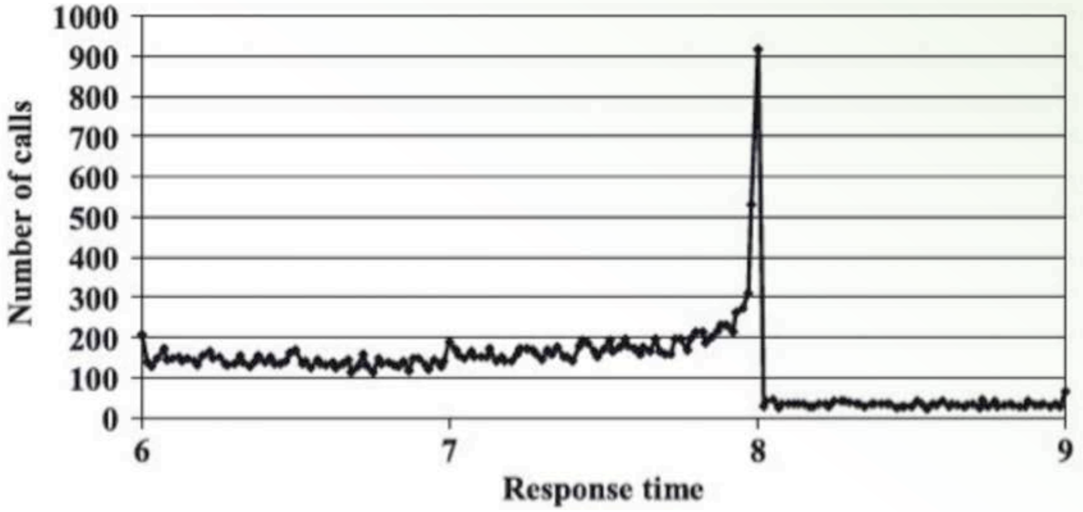
Figure 3. Percentage of category A calls met within eight minutes, in England (■), Wales (●) and Scotland (△)



Sources: England, Department of Health (1999a, 2000, 2001) (for 1999–2001) and Information Centre (2007) (for 2002–07); Wales, National Assembly for Wales (2005) (for 2000–04), Auditor General for Wales (2006), page 37 (for 2005 and 2006), and Welsh Assembly Government (2007b) (for 2007); Scotland, Auditor General for Wales (2006), page 37, and Audit Scotland (2007), page 2 (for 2007). Bevan and Hamblin (2009).

Ambulance trusts gamed – instead of the 'noisy' decline in response times one would expect to see, there were clear jumps and discontinuities in numbers of calls reported answered just before and after the eight-minute mark, where longer response times were reassigned under the eight-minute mark, for example (Figure 4).[82]

Figure 4. Spike in frequency of response times to category A calls at eight minutes by one service



Source: Commission for Health Improvement (2003c); Bevan and Hamblin (2009).

However, subsequent analysis showed ambulance trusts were 'gilding the lily' of what they'd already achieved by legitimate means – such gaming had only improved an already better score.[82]

'Wherever performance or quality are measured, then there will be an incentive to manipulate the data or circumstances to give the appearance of compliance,' says Mears (2014).[81] Indeed, Mears argues that the existence of gaming behaviour is in fact an indication of a successful performance measurement scheme being taken seriously: 'Organisations will only engage in gaming behaviour where the incentive is powerful and the measure effective in discerning good from poor performance.'[81] Further, say Bevan and Hamblin (2009), 'If we do take systems of performance measurement seriously, and design these to have an effect, then developing systems to counter gaming ought to be integral to the design of such systems.'[82]

The negative unintended consequence often cited of such ratings are the damage to morale for those working in underperforming institutions. However, Bevan and Hamblin suggest it 'can be argued that damaging morale is necessary in the short term for creating the different atmosphere that is required to achieve improvement in the long term'.[82]

Contandriopolous et al (2014), in analysing the causal pathways between public reporting mechanisms and effects on quality, support Hibbard et al: 'There is convincing evidence that [public reporting of performance measures] triggers effects (such as organizational efforts to improve performance) that private disclosure of the same data to organizational management fails to produce.'[48, 56, 61, 146, 147]

That is, public reporting of quality measures, properly done, makes health care better.

Systematic review and evidence report

A systematic review by Fung et al (2008) of public reporting of performance measurement and its effects on quality improvement found that public reporting of performance data stimulates quality improvement at the hospital level.[125] The review looked at 45 articles assessing public release of performance data (including patient experience surveys) on selection of providers, quality improvement activity, clinical outcomes and unintended consequences. The evidence strength is moderate, but suggests publicly releasing performance data stimulates quality improvement at the hospital level but not at an individual level. There is, however, equivocal evidence that contradicts this finding, in relation to the American College of Surgeons National Surgical Quality Improvement Program surgical outcomes public reporting initiative, which found voluntary hospital reporting was not associated with better performance on 54 of 58 measures.[148]

The large 2012 evidence report produced by the US Agency for Healthcare Research and Quality found that, in general, 'public reporting is more likely to be associated with changes in health care provider behavior than with selection of health services providers by patients or families'.[55] Providers engage in activities to improve quality of care when their performance data are made public: 'Quality measures that are publicly reported improve over time.'[55]

When the evidence report focused on public reporting of data on the performance of individual clinicians, the literature was limited. Most of the review's quantitative studies on individual reporting were for cardiac surgery (10 of 12), and eight of those studies were on the NYS CSRS (see below).[54, 63, 149-154] Two were from the Pennsylvania Guide to Coronary Artery Bypass Graft Surgery.[155, 156] The researchers found moderate evidence that individual clinicians 'respond to public reporting in positive ways, including adding services, changing policy, and increasing focus on clinical care. One study found that low-quality surgeons leave practice (considered a positive action).' High-risk patients were still seen to have high-quality surgeons, 'which is counter to the hypothesis that public reporting might cause adverse selection'.[149] 'Surgeon-specific mortality rates for CABG in New York State declined after rates were publicly reported (one study).'[63]
See below for more information on the nature of the NYS CSRS.

In this review the data on public reporting on individual clinicians are limited predominantly to cardiac surgery, and mostly relate to the NYS CSRS. Qualitative studies (20 were included) suggested surgeons and clinicians had fears that were not borne out – that they would not operate on high-risk patients, for example.[157] A UK survey conducted in 2005 and 2009 found that though cardiac surgeons still did not welcome reporting of individual surgeons' results, the percentage declined over time (68.8 percent in 2005; 43.4 percent in 2009) and an increasing percentage believed such reporting actually improved standards (42.2 percent in 2005; 64.9 percent in 2009).[158]

New York State Cardiac Surgery Reporting System (NYS CSRS)

The NYS CSRS is the first and longest-running statewide programme in the US to produce data on outcomes for cardiac surgery.[62] It is also the most studied. The scheme was put in place at the directive of the State Commissioner for Health, who in 1988 became alarmed at a fivefold variation in New York hospital mortality rates for CABG procedures.[10] Their initial work determined that administrative data were insufficient and a patient-level clinical registry model was pursued. Later studies would show that the two sources produced different conclusions about relative quality of care, and that the registry data were more predictive of mortality – a relevant consideration in the New Zealand context where predominantly administrative data are available at present.[159, 160]

The initial New York risk adjustment model incorporated patient risk factors from the literature at the time 'with demographics, complications of care, admission and discharge dates, procedures performed, and patient disposition at discharge'.[10] The observed, expected, and risk-adjusted mortality rates and volumes by unnamed hospital were first published in 1990, by which time hospitals had been receiving their own data confidentially since the first half of 1989.[161] Mortality rates for the confidential release period, the first half of 1990, had declined by 14 percent from the first half of 1989. The same day the anonymous hospital data were published, the Commissioner released the names of the hospitals to *The New York Times*. The newspaper *Newsday* then sued the Department of Health to release surgeon-specific outcome information, and prevailed.[10]

The trajectory is relevant to the New Zealand context – however, the registry existed prior to the release and was tasked with assessing the relative quality of cardiac care.

In December 1992 the NYS CSRS released surgeon-specific outcome data on a rolling three-year basis only for surgeons with greater than 200 cases within that time frame, in order to generate greater statistical power. Hannan et al (2012) describe quality improvement activities pursued by 'several' hospitals in response to data showing their having greater than expected risk-adjusted mortality rates.[10] These activities differed by hospital, and were arrived at via internal investigation and review spurred by the data, and included improved management of stabilisation before surgery,[162] suspension and review of entire cardiac services,[62] and establishment of quality assurance programmes, credentialing programmes, and dedicated facilities and staff.[62] These activities were associated with varying degrees of decreased mortality rates for these hospitals.[10, 162]

The New York programme has been associated with a 41 percent statewide fall in in-hospital mortality rates for CABG surgery from 4.17 percent in 1989 to 2.45 percent in 1992.[10, 63] Attribution and accuracy of this figure has been a matter of controversy, and challenges have been mounted on the basis of outmigration of sicker patients,[64] reluctance to treat sicker patients, and gaming, including upcoding of sicker patients to increase casemix complexity and the expected number of deaths (reducing the observed/expected mortality index).[45, 65-67]

Further, increased sophistication of risk adjustment analysis (including so-called 'reliability adjustment') to decrease statistical noise from hospital quality rankings was found to greatly reduce variation between hospital mortality rates for patients undergoing colon resection in 2007 (n = 181 hospitals, n = 18,455 patients).[163] The extent of the effects of these techniques on the New York findings are unknown, but the picture overall is equivocal.[66]

California Cardiac Surgery and Intervention Project (CCSIP)

In the case of mortality outcomes after cardiac procedures in New York, carefully considered public reporting of individual surgeons' outcomes appeared to act as a spur to improve quality. There are now several similar statewide efforts.[68-71] In California, the introduction of the CCSIP, similar to the NYS CSRS, 'was associated with improved outcomes'.[72] The CCSIP has an explicit quality improvement focus, describing the programme as 'dedicated to data collection, study of best practices, and performance improvement since 2003'.[164]

A study of the CCSIP found that after its introduction in 2003, 'risk-adjusted in-hospital mortality for CABG, CABG plus valve or aneurysm, and valve procedures decreased during 2003 and 2004 compared with 1998 through 2002, and PCI [percutaneous coronary intervention] mortality remained unchanged'.[72] Five years later a follow-up study found that in the Californian context the public release of hospital performance reports on CABG procedures showed at the hospital level no indication of reduced risk-adjusted mortality overall, and 'was associated with increased volume at low-mortality hospitals, and may have reduced referrals of high-risk patients to high-mortality hospitals (or risk avoidance)'.[73] Like the NYS CSRS, the CCSIP had a highly evolved risk-adjustment process and a focus on improvement and on clinician buy-in. The common outcome whether publicly reported or not is the importance of structured initiatives to collect high-quality clinical data, to use the data to determine risk-adjusted outcome rates and feedback of results to providers.

The UK trajectory

Public reporting of individual surgeons' outcomes in the UK had its roots in the paediatric cardiac surgery scandal at Bristol Royal Infirmary, and the subsequent investigation led by Sir Ian Kennedy.[75] Information and how it is used formed a chapter of the report, and generated several recommendations. Recommendation

155 of the report stated, ‘Patients and the public must be able to obtain information as to the relative performance of the trust and the services and consultant units within the trust.’ This is actually a rather nuanced recommendation – there is no mention of mortality rates and the expression ‘consultant units’ is synonymous with team rather than individual surgeon.

The Department of Health responded positively to this suggestion: ‘We wrote with the support of the BMA, to NHS consultants in December 2001 telling them of our intentions to use available data to publish performance information at consultant team level.’[21]

The SCTS was comparatively well placed to respond to this having had a voluntary comprehensive audit data set in place since 1996,[76] although the issues of risk adjustment were complex, even allowing for good data.[58]

In 2001, the Dr Foster organisation published mortality for coronary artery surgery for named cardiac surgical hospitals throughout the UK. In March 2005, named surgeon mortality data for coronary artery surgery for all UK surgeons were published by *The Guardian* newspaper, after a request under the newly introduced UK Freedom of Information Act. In April 2006 a website was launched by the Healthcare Commission and the SCTS, which provided cardiac surgery mortality data for the public. Ownership of the shared website reverted entirely to the SCTS in 2010 – and this has been published since as the ‘Blue Book Online’.[77]

Outcomes have clearly improved in the NHS[77–79] in the period since publication, and casemix has become more complex, which provides evidence that ‘despite clinical concerns, the introduction of public accountability has not led to a decrease in the number of high risk patients coming for coronary artery surgery’.[78] A causal link from publication to reduced mortality has not been shown.

However, there remains angst among the cardiac surgeons a decade after first publication. The SCTS wrote to the NHS England CEO in January 2015 asking for a reconsideration of publication, precisely because of the concern of risk selection:

> This debate is primarily driven by concern that publishing at consultant level results in risk-averse behaviour, with some surgeons less willing to operate on patients with a perceived high risk of mortality, despite potential overall patient benefit, because of a fear for the potential adverse consequences for themselves if the patient dies.[80]

Clinical registries for quality improvement

A clinical registry is a structured ongoing collection of personal health data arising from all patients in a clinically defined population, for the purpose of systematically and continuously improving the quality and safety of health care for that population. Once established, registries can provide rapid feedback concerning variations in processes and outcomes of care, benchmark these across providers, and look at appropriateness, access, safety and effectiveness of care. A key strength is the ability to risk-adjust outcomes for severity, comorbidities and other covariates.[19]

Clinical registries began as a data source used by health care practitioners. Although still clinically led, they have now become an important quality improvement tool, and can provide feedback to different levels of the health care system, including individual practitioners, health care providers and central agencies, as well as consumers. Some countries such as Sweden have invested in this approach with over 100 clinical quality registries. Registries can be a powerful tool for clinical quality improvement, with the biggest impact coming from feeding back to units their performance in comparison with their peers. This feedback, along with the availability of best practice guidelines to guide improvement, relies on strong clinical governance processes to bring about the necessary quality improvements. In New Zealand, Health Roundtable data are provided to DHBs in a similar way, though the feedback is based on DHB administrative data rather than registries.

Should New Zealand be developing clinical registries for quality improvement?

Clinical quality registries tend to be of most value in areas of clinical practice where an intervention or pattern of care is discrete and well defined, such as surgical or other invasive clinical interventions, acute episodes of care such as admissions to an intensive care unit, discrete diseases (eg, multiple sclerosis), and discretely defined chronic disease management (eg, renal dialysis). They are less useful in conditions (such as hypertension or diabetes) where the borderline between who is and is not eligible for inclusion is unclear.[19]

Registries have traditionally been established as freestanding databases, with data collected through paper or online forms. However, there is an increasing focus on capturing the data electronically during the routine process of clinical care and subsequently transferring the data to a registry database. Currently registries are expensive (as mentioned, the New Zealand cardiac registry costs approximately $1 million per year – K Evison, Ministry of Health, personal communication, 3 July 2015) as our IT systems are not configured to electronically capture the required data as part of day-to-day clinical practice. Thus the case for investing in IT systems that enable automatic data capture as part of routine care needs to be considered.

Even if our system allowed such capture in developing registries, clinical priority areas need to be identified, various parties brought together, a governance structure developed, the data model and definitions designed, and a repeated cycle of data collection, bio-statistical analysis, clinical interpretation, and feedback to providers established. It then requires the clinical governance structures locally to identify and implement changes to improve health care quality. Bridgewater et al (2013) describe the establishment and governance of the UK National Adult Cardiac Surgery Audit.[4] This programme has driven a culture of data collection among surgeons, and now measures a variety of postoperative outcomes including new renal intervention, stroke, and re-exploration for bleeding in addition to mortality.

A limitation of registries is that patient-reported outcomes are not part of the process but are collected from patients at defined intervals after the baseline episode of care. These are often done through telephone or online surveys so would need to continue to be separately captured and reported alongside registries.

New Zealand currently has some established clinical registries – for example, the cardiac registry and joint registries. Some sit within the Ministry of Health; others within the oversight of colleges. Some, like the cardiac registry, have been designed with clinician and consumer input with plans for publication. Currently there is no consistent approach to registries in New Zealand but we need to consider their role as part of this overall process.

In the UK, bridging of the administrative and clinical collection of data was highly problematic. Recommendation 148 in the Bristol Royal Infirmary Inquiry ('the Kennedy report') stated:

> The current 'dual' system of collecting data in the NHS in separate administrative and multiple clinical systems is wasteful and anachronistic. A single approach to collecting data should be adopted, which clinicians can trust and use and from which information about both clinical and administrative performance can be derived.[75]

The UK Department of Health replied:

> We agree. Those responsible for the separate administrative and clinical audit databases are already working together to develop an approach which will avoid duplication. Implementation of *Information for Health* will provide the basis for a single approach to collecting data for both clinical and administrative needs through the electronic patient record, which will be introduced by 2005.[21]

Unfortunately, implementation of the NHS National Programme for IT has entered the annals of failed large-scale IT projects – the reality is more complex than the theory. The system has been axed, though portions of it remain in place.[22]

TEAMWORK: WHAT EVIDENCE IS THERE THAT INDIVIDUALS ARE SOLELY RESPONSIBLE FOR A PATIENT'S OUTCOME?

'Measure your team's performance, and publish the results,' wrote the editor of the *BMJ* in 2012. 'But where is the clinical leadership pushing for public access to performance data of individual clinical teams?'[15, 165]

The drive from the colleges and professional bodies for greater transparency aside, reporting by individual surgeon ignores the effect of the multidisciplinary team and the context in which especially complex surgery is done. Many aspects of care other than the surgeon's performance will affect the outcome, such as timeliness of referral and diagnosis, perioperative and postoperative care, and follow-up care after discharge.[27]

The enquiry into deaths of paediatric patients undergoing cardiac surgery at Bristol Royal Infirmary revealed lack of leadership, teamwork and the ability to work together effectively for the interests of patients at the hospital.[61] Although the surgeon plays an important role, so does the anaesthetist, the intensive care physician, the intensive care nurse, the socioeconomic status of the local population, the severity of cardiac illness, comorbidities, adequacy of facilities and staffing levels, attitude to training, interpersonal relationships between staff, and geographical layout of the unit, to name some of the other considerations at play.[61]

Such findings have led to the decision by the US Veterans Health Administration to discourage surgeon-specific outcomes as they believe the performance cannot be separated from that of the institution.[61] Strong support for outcomes not being directly attributable to an individual was shown by Huckman and Pisano in 2006. The researchers found that cardiac surgical patient mortality rates did not follow a particular surgeon moving between institutions. Though an increased volume of procedures performed by the surgeon at one hospital led to improved mortality rates there, an increased volume of procedures at another hospital did not. Their performance was not fully 'portable'. This 'firm specificity' showed clearly that patient outcomes were not tied to an individual surgeon; they were dependent on other factors: team, facility and organisation.[31] Teamwork and communication are crucial to reducing error.[172-174]

Others argue that doctors are best placed to change institutional processes that affect outcomes and so therefore are a logical target.[61]

The 2012 Agency for Healthcare Research and Quality systematic review concurs with the team argument: 'Approaches to health care that are anchored in team's and systems' responses to assure safety are contrary to the idea that any one individual is solely responsible for outcomes.'[55]

In their 2015 paper, Wong et al looked at low- and high-mortality hospitals doing cancer surgery, all with sufficient caseloads, and examined both medical and surgical complications.[85] Rates of complications were not significantly different between the two groups but case fatality was significantly higher. Drilling down, specific procedural complication rates were similar but again the case fatality was greater in the high mortality hospitals. The difference emerged in the so-called 'failure to rescue' rates. Although complication rates were indistinguishable, case fatality was approximately twice as high. The paper then discusses contributing factors such as better management of multisystem organ failure in intensive care units, hospital resources such as staffing levels and rapid response teams, and organisational factors of safety culture, teamwork and communication.[85]

McCrum et al (2013) showed that top-performing hospitals had lower overall hospital mortality rates even when risk-adjusted. They suggested that understanding the systems and leadership characteristics could help identify components of a truly good hospital that can be used to improve mortality rates at the lower-performing institutions.[166] Organisational culture was raised in feedback from the Public Service Association: a high-trust workplace culture was needed in New Zealand such that information was gathered from reflective practice in a no-blame environment so learning and experience could be safely shared.

Westaby et al (2015) proposed that named surgeon reporting undermines the importance of teamwork and may divert attention from key aspects of process and management that might impact directly on mortality rates.[167]

The authors looked at surgeon-specific mortality data and failings in delivery of safe surgical services and analysed deaths to identify the root cause that was the catalyst for a cascade of processes culminating in death. The study looked at 1500 consecutive patients operated on by three surgeons in the UK with autopsies done in 86 percent of cases. The cases were independently assessed by a panel and broken down into three classes:

- class I – surgeon-dependent related to the surgical procedure, either inappropriate selection, an intraoperative event, a technical error or post-op bleeding
- class II – failure to rescue from common conditions such as renal failure, pneumonia and cardiac arrests
- class III – multifactorial including frailty, critical preoperative state or surgical complexity.

Long operating times were deemed surgeon-related. Of the 51 deaths none occurred in the operating theatre and none of the autopsies revealed a surgical error as a precipitating factor. There were 37 percent in class I, 43 percent in class II, and the rest class III. The median age of the 51 who died was 11 years older at 79, and their expected mortality was greater in comparison to the overall group. There were higher numbers of females in all the categories and class I mortality was principally related to urgent status. Failures to rescue rates vary between organisations. Team consistency, staffing levels, nurse education, job satisfaction and burnout have all been identified as factors that underpin failure to rescue. Studies have shown that each additional patient per nurse was associated with a 7 percent increase in the odds of failure to rescue, which was therefore related to effective human resource management and hospital systemic factors. A number of deaths were out-of-hours events which involved inexperienced temporary staff.[167] Studies show higher mortality for complex surgery at the end of the working week or weekend.[168]

Neily et al (2010) compared Veterans Health Administration facilities where surgical/operating room team training was given (using crew resource management theories adapted from aviation) with facilities where such training had not been given. Over a day clinicians were trained to work as a team, challenge each other when they identified safety risks, conduct checklists pre-op and post-op briefings and debriefings, and use strategies such as red flags and communication during care transitions.[90] The 74 trained facilities showed a significant decrease of 18 percent in observed mortality versus 7 percent in the non-trained. For each quarter of training the mortality rate decreased by 0.5 per 1000 procedure deaths. Armour Forse (2011) showed that operating room team training improved overall surgical morbidity and mortality, which were both significantly improved (mortality, 2.7 percent to 1 percent ($P < .05$); morbidity, 20.2 percent to 11.0 percent ($P < .05$)) with a need to continue team training to provide sustained improved operating room culture.[169]

In a 2014 systematic review examining studies of team training interventions published between 2000 and 2012, Weaver et al showed 13 studies reported statistically significant changes in teamwork behaviour processes and 10 reported improvements in clinical care processes or patient outcomes including mortality and morbidity.[170] Effects were noted across a range of different clinical contexts but larger impacts were reported for interventions that included tools and organisational changes to support and sustain teamwork competencies into daily practice. The paper concluded there is moderate- to high-quality evidence suggesting team training can impact on processes and patient outcomes.[170] Shekelle et al (2013) identify team training as a top patient safety strategy for adoption.[107]

Merry et al (2014) argue that outcomes for cardiac surgical patients depend on the collaborative effort of the whole team of health practitioners involved in the patient's care, from decisions on treatment plans, through to pre-operative, intra-operative and post-operative care and discharge planning.[171] Combining the data and team arguments, Burger et al (2007) make the argument that, due to low statistical power, an individual surgeon's risk estimate may be less accurate for that surgeon than a weighted average of data from their own performance, their colleagues, and their institution.[27]

If outcomes are not therefore dependent on the individual but the team and organisational factors, accountability to the public would appear to require us to publish at team or organisation level as these factors affect outcomes more than the individual surgeon. Further, at a team level there may be sufficient data to reach meaningful statistical power and allow a timelier pick-up of issues. The sector also needs to consider, as did the

New Zealand consumer workshop group, what organisational factors – such as safety culture; staffing levels and education; teamwork and team education; and rapid response teams – should be made available as these are key to good outcomes. The evidence would appear to point to team and organisational systemic factors as the real measures to provide.

References

1. Keogh B. 2008. Presidential Message: A decade of immunisation. Society for Cardiothoracic Surgery in Great Britain & Ireland.
URL: www.scts.org/_userfiles/pages/file/Presidential%20address-Keogh.pdf (accessed September 2015).
2. Chaudhry M, Gauld R, Horsburgh S. 2012. Hospital quality-of-care performance measurement and reporting: what New Zealand can learn from the United States and United Kingdom. *NZ Med J* 125(1366): 51–63.
3. Davies H, Marshall MN. 1999. Public disclosure of performance data: does the public get what the public wants? *The Lancet* 353(9165): 1639–40.
4. Bridgewater B, Hickey GL, Cooper G, et al. 2013. Publishing cardiac surgery mortality rates: lessons for other specialties. *BMJ* 346(feb28): f1139.
5. Health Quality and Safety Commission; Ministry of Health. Consumer workshop. 3 July 2015.
6. Sinaiko AD, Eastman D, Rosenthal MB. 2012. How report cards on physicians, physician groups, and hospitals can have greater impact on consumer choices. *Health Affairs* 31(3): 602–11.
7. Minami CA, Dahlke A, Bilimoria KY. 2015. Public Reporting in Surgery. *Annals of Surgery* 261(2): 241–2.
8. Chen J, Ou L, Hollis SJ. 2013. A systematic review of the impact of routine collection of patient reported outcome measures on patients, providers and health organisations in an oncologic setting. *BMC Health Services Research* 13(1): 211.
9. Greenhalgh J, Long AF, Flynn R. 2005. The use of patient reported outcome measures in routine clinical practice: lack of impact or lack of theory? *Soc Sci Med* 60(4): 833–43.
10. Hannan EL, Cozzens K, King SB 3rd, et al. 2012. The New York State cardiac registries: history, contributions, limitations, and lessons for future efforts to assess and publicly report healthcare outcomes. *J Am Coll Cardiol* 59(25): 2309–16.
11. Morsi E, Lindenauer PK, Rothberg MB. 2012. Primary care physicians' use of publicly reported quality data in hospital referral decisions. *J Hosp Med* 7(5): 370–5.
12. Shahian DM, Edwards FH, Jacobs JP, et al. 2011. Public Reporting of Cardiac Surgery Performance: Part 1–History, Rationale, Consequences. *Ann Thorac Surg* 92(3): S2–S11.
13. Neil DA, Clarke S, Oakley J. 2007. *Against the Informed Consent Argument for Surgeon Report Cards. Informed Consent and Clinician Accountability.* Cambridge: Cambridge University Press. p 147–56.
14. Walker K, Neuburger J, Groene O, et al. 2013. Public reporting of surgeon outcomes: low numbers of procedures lead to false complacency. *The Lancet* 382(9905): 1674–7.
15. Godlee F. 2012. Measure your team's performance, and publish the results. *BMJ* 345: e4590.
16. Banja JD. 2015. Disclosure of experience as a risk factor in informed consent for neurosurgery: the case of Johnson v. Kokemoor. *Virtual Mentor* 17(1): 69–73.
17. Birkmeyer JD, Stukel TA, Siewers AE, et al. 2003. Surgeon volume and operative mortality in the United States. *N Engl J Med* 349(22): 2117–27.
18. Hannan EL, Siu AL, Kumar D, et al. 1995. The decline in coronary artery bypass graft surgery mortality in New York State. The role of surgeon volume. *JAMA* 273(3): 209–13.
19. University of Melbourne. 2013. *Registry Science Handbook.* Melbourne: University of Melbourne.
20. Bridgewater B, Livesey S, Kinsman R. 2011. *Maintaining patients' trust: modern medical professionalism.* Society for Cardiothoracic Surgery in Great Britain and Ireland.
21. Department of Health. *Learning from Bristol: The Department of Health's Response to the Report of the Public Inquiry into children's heart surgery at the Bristol Royal Infirmary 1984–1995.* UK: Department of Health.

22. Maughan A. 2010. Six reasons why the NHS National Programme for IT failed. *ComputerWeekly.com.* URL: http://www.computerweekly.com/opinion/Six-reasons-why-the-NHS-National-Programme-for-IT-failed (accessed September 2015).
23. Johnson v Kokemoor, *199 Wis2d 615, 545 NW2d 495.* 1996. Supreme Court of Wisconsin.
24. *Avila v Flangas, No. 04-95-00106-CV, 1996 WL 63036, at *1 (Tex App).* 1996. Texas Court of Appeals.
25. *Duttry v. Patterson, 741 A.2d 199.* 1999. Supreme Court of Pennsylvania.
26. Bal BS, Choma TJ. 2012. What to disclose? Revisiting informed consent. *Clin Orthop Relat Res* 470(5): 1346–56.
27. Burger I, Schill K, Goodman S. 2007. Disclosure of individual surgeon's performance rates during informed consent: ethical and epistemological considerations. *Annals of Surgery* 245(4): 507–13.
28. Edozien LC. 2015. UK law on consent finally embraces the prudent patient standard. *BMJ* 350: h2877.
29. Sokol DK. 2015. Update on the UK law on consent. *BMJ* 350: h1481.
30. Governor-General. 1996. *Health and Disability Commissioner (Code of Health and Disability Services Consumers' Rights) Regulations (SR 1996/78).* Wellington: Governor-General.
31. Huckman RS, Pisano GP. 2006. The firm specificity of individual performance: evidence from cardiac surgery. *Management Science* 52(4): 473–88.
32. Lagu T, Goff SL, Hannon NS, et al. 2013. A mixed-methods analysis of patient reviews of hospital care in England: implications for public reporting of health care quality data in the United States. *Jt Comm J Qual Patient Saf* 39(1): 7–15.
33. Henderson A, Henderson S. Provision of a surgeon's performance data for people considering elective surgery. *Cochrane Database Syst Rev* 2015, Issue 2, Art. No. CD006327.
34. Ketelaar NA, Faber MJ, Flottorp S, et al. 2011. Public release of performance data in changing the behaviour of healthcare consumers, professionals or organisations. *Cochrane Database Syst Rev 2011, Issue 11, Art. No.* CD004538.
35. Alexander JA, Hearld LR, Hasnain-Wynia R, et al. Consumer Trust in Sources of Physician Quality Information. *Medical Care Research and Review* 68(4): 421–40.
36. Schneider EC, Epstein AM. 1998. Use of public performance reports: a survey of patients undergoing cardiac surgery. *JAMA* 279(20): 1638–42.
37. Mannion R, Goddard M. 2003. Public disclosure of comparative clinical performance data: lessons from the Scottish experience. *J Eval Clin Pract* 9(2): 277–86.
38. Donelan K, Rogers RS, Eisenhauer A, et al. 2011. Consumer Comprehension of Surgeon Performance Data for Coronary Bypass Procedures. *Ann Thorac Surg* 91(5): 1400–06.
39. Faber M, Bosch M, Wollersheim H, et al. 2009. Public reporting in health care: how do consumers use quality-of-care information? *Medical Care* 47(1): 1-8.
40. Robinowitz DL, Dudley RA. 2006. Public reporting of provider performance: can its impact be made greater? *Ann Rev Public Health* 27(1): 517–36.
41. Marshall MN. 2004. How do we maximize the impact of the public reporting of quality of care? *Int J Qual Health Care* 16(suppl_1): i57–i63.
42. Marshall MN, Shekelle PG, Leatherman S, et al. 2000. The public release of performance data: what do we expect to gain? A review of the evidence. *JAMA* 283(14): 1866–74.
43. Seddon M. 2012. Reporting hospital performance – a balancing act between accountability and quality improvement. *NZ Med J* 125(1366): 5–8.
44. Schlesinger M, Kanouse DE, Martino SC, et al. 2013. Complexity, public reporting, and choice of doctors: a look inside the blackest box of consumer behavior. *Med Care Res Rev* 71(5 Suppl): 38S–64S.

45. Hamblin R. 2007. Measurements, incentives, and improvement: what are the advantages, what are the pitfalls and what might work? A review of incentivized measurement schemes in the US and UK since 1990. Unpublished manuscript.

46. Huckman RS, Kelley MA. 2013. Public reporting, consumerism, and patient empowerment. *N Engl J Med* 369(20): 1875–7.

47. Hussey PS, Luft HS, McNamara P. 2014. Public reporting of provider performance at a crossroads in the United States: summary of current barriers and recommendations on how to move forward. *Med Care Res Rev* 71(5 Suppl): 5S–16S.

48. Hibbard JH, Stockard J, Tusler M. 2005. Hospital performance reports: impact on quality, market share, and reputation. *Health Affairs* 24(4): 1150–60.

49. Christianson J, Maeng D, Abraham J, et al. 2014. What influences the awareness of physician quality information? Implications for Medicare. *Medicare & Medicaid Research Review* 4(2): E1–E15.

50. Davies HTO. 2002. Health care report cards: implications for vulnerable patient groups and the organizations providing them care. *J Health Polit Policy Law* 27(3): 379–400.

51. Dimick J, Ruhter J, Sarrazin MV, et al. 2013. Black patients more likely than whites to undergo surgery at low-quality hospitals in segregated regions. *Health Affairs* 32(6): 1046–53.

52. Fung CH. 2008. Systematic review: the evidence that publishing patient care performance data improves quality of care. *Ann Intern Med* 148(2): 111.

53. SteelFisher GK, Schneider EC, Zaslavsky AM, et al. 2009. Can quality reports help address health care disparities? Use and awareness of comparative quality information by African Americans. *J Health Care Poor Underserved* 20(3): 884–95.

54. Werner RM. 2005. Racial profiling: the unintended consequences of coronary artery bypass graft report cards. *Circulation* 111(10): 1257–63.

55. Totten AM, Wagner J, Tiwari A, et al. 2012. Closing the quality gap: revisiting the state of the science (vol. 5: public reporting as a quality improvement strategy). *Evid Rep Technol Assess (Full Rep)* (208.5): 1–645.

56. Contandriopoulos D, Champagne F, Denis JL. 2014. The multiple causal pathways between performance measures' use and effects. *Med Care Res Rev* 71(1): 3–20.

57. Berwick DM, James B, Coye MJ. 2003. Connections between quality measurement and improvement. *Med Care* 41(1 Suppl): I30–8.

58. Bevan G, Hood C. 2006. Have targets improved performance in the English NHS? *BMJ* 332(7538): 419–22.

59. Hamblin R. 2007. Publishing 'quality' measures: how it works and when it does not? *Int J Qual Health Care* 19(4): 183–6.

60. Hamblin R. 2008. Regulation, measurements and incentives. The experience in the US and UK: does context matter? *J R Soc Promot Health* 128(6): 291–8.

61. Hibbard JH, Stockard J, Tusler M. 2003. Does publicizing hospital performance stimulate quality improvement efforts? *Health Aff (Millwood)* 22(2): 84–94.

62. Chassin MR. 2002. Achieving and sustaining improved quality: lessons from New York State and cardiac surgery. *Health Aff (Millwood)* 21(4): 40–51.

63. Hannan EL, Kumar D, Racz M, et al. 1994. New York state's cardiac surgery reporting system: Four years later. *Ann Thorac Surg* 58(6): 1852–7.

64. Omoigui NA, Miller DP, Brown KJ, et al. 1996. Outmigration for coronary bypass surgery in an era of public dissemination of clinical outcomes. *Circulation* 93(1): 27–33.

65. Burack JH, Impellizzeri P, Home P, et al. 1999. Public reporting of surgical mortality: a survey of New York State cardiothoracic surgeons. *Ann Thorac Surg* 68(4): 1195–200.

66. Epstein AJ. 2006. Do cardiac surgery report cards reduce mortality? Assessing the evidence. *Med Care Res Rev* 63(4): 403–26.

67. Narins CR, Dozier AM, Ling FS, et al. 2005. The influence of public reporting of outcome data on medical decision making by physicians. *Arch Intern Med* 165(1): 83.

68. Harvard Medical School. 2011. *Adult coronary artery bypass graft surgery in the Commonwealth of Massachusetts: fiscal year 2009 report.* Boston: Harvard Medical School.

69. California's Office of Statewide Health Planning and Development. 2011. *The California report on coronary artery bypass graft surgery: 2007 hospital data.* California: OSHPD.

70. New Jersey Department of Health and Senior Services, Office of Health Care Quality Assessment. 2011. *Cardiac surgery in New Jersey: 2008.* Trenton: New Jersey Department of Health and Senior Services.

71. Pennsylvania Health Care Cost Containment Council. 2010. *Cardiac surgery in Pennsylvania 2007–2008.* Harrisburg: Pennsylvania Health Care Cost Containment Council.

72. Carey JS, Danielsen B, Junod FL, et al. 2006. The California Cardiac Surgery and Intervention Project: evolution of a public reporting program. *Am Surg* 72(10): 978–83.

73. Romano PS, Marcin JP, Dai JJ, et al. 2011. Impact of public reporting of coronary artery bypass graft surgery performance data on market share, mortality, and patient selection. *Medical Care* 49(12): 1118–25.

74. Weiser TG, Semel ME, Simon AE, et al. 2011. In-hospital death following inpatient surgical procedures in the United States, 1996-2006. *World J Surg* 35(9): 1950–6.

75. Kennedy I. 2001. *The report of the public inquiry into children's heart surgery at the Bristol Royal Infirmary 1984-1995: learning from Bristol.* ('The Kennedy report'.)
URL: http://webarchive.nationalarchives.gov.uk/20090811143745/http:/www.bristol-inquiry.org.uk/final_report/the_report.pdf (accessed September 2015).

76. Keogh B, Kinsman R. 2002. *National adult cardiac surgical database report 2000-2001.* Society of Cardiothoracic Surgeons of Great Britain and Ireland. London: Society of Cardiothoracic Surgeon of Great Britain and Ireland.

77. Society for Cardiothoracic Surgery in Great Britain and Ireland (SCTS). *Blue Book Online.* 2015.
URL: http://bluebook.scts.org/ (accessed September 21).

78. Bridgewater B, Grayson AD, Brooks N, et al. 2007. 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* 93(6): 744–8.

79. Brown KL, Crowe S, Franklin R, et al. 2015. Trends in 30-day mortality rate and case mix for paediatric cardiac surgery in the UK between 2000 and 2010. *Open Heart* 2(1): e000157–e57.

80. Boseley S. 2015, 30 January. Surgeons ask NHS England to rethink policy of publishing patients' death rates. *Guardian.*
URL: http://www.theguardian.com/society/2015/jan/30/surgeons-nhs-england-patients-death-rates-bruce-keogh-jeremy-hunt-health (accessed September 2015).

81. Mears A. 2014. Gaming and targets in the English NHS. *Universal Journal of Management* 2(7): 293–301.

82. Bevan G, Hamblin R. 2009. Hitting and missing targets by ambulance services for emergency calls: effects of different systems of performance measurement within the UK. *J R Stat Soc Ser A Stat Soc* 172(1): 161–90.

83. Joint Commission. 2006. *Joint Commission 2006 Sentinel Event Statistics 1995–2005.* Washington: Joint Commission.

84. Wolf FA, Way LW, Stewart L. 2010. The efficacy of medical team training: improved team performance and decreased operating room delays: a detailed analysis of 4863 cases. *Ann Surg* 252(3): 477–83; discussion 83-5.

85. Wong SL, Revels SL, Yin H, et al. 2015. Variation in hospital mortality rates with inpatient cancer surgery. *Annals of Surgery* 261(4): 632–6.

86. Glance LG, Kellermann AL, Hannan EL, et al. 2015. The impact of anesthesiologists on coronary artery bypass graft surgery outcomes. *Anesth Analg* 120(3): 526–33.

87. Leslie K, Merry AF. 2015. Cardiac surgery: all for one and one for all. *Anesth Analg* 120(3): 504–6.

88. Merry AF, Ramage MC, Whitlock RM, et al. 1992. First-time coronary artery bypass grafting: the anaesthetist as a risk factor. *Br J Anaesth* 68(1): 6–12.

89. Slogoff S, Keats AS. 1985. Does perioperative myocardial ischemia lead to postoperative myocardial infarction? *Anesthesiology* 62(2): 107–14.

90. Neily J, Mills PD, Young-Xu Y, et al. 2010. Association between implementation of a medical team training program and surgical mortality. *JAMA* 304(15): 1693.

91. Keogh B. 2004. The legacy of Bristol: public disclosure of individual surgeons' results. *BMJ* 329(7463): 450–4.

92. Berwick D. 2013. *A promise to learn – a commitment to act: improving the safety of patients in England.* National Advisory Group on the Safety of Patients in England. UK Department of Health. URL: https://www.gov.uk/government/publications/berwick-review-into-patient-safety (accessed September 2015.)

93. Francis R. 2013. *The Mid Staffordshire NHS Foundation Trust Public Inquiry.* URL: http://webarchive.nationalarchives.gov.uk/20150407084003/http://www.midstaffspublicinquiry.com/report (accessed September 2015). London: The Mid Staffordshire NHS Foundation Trust Public Inquiry.

94. Roberts RG, Friedsam D, Beasley JW, et al. 2006. The state of quality reports: comparing states by their rankings. *WMJ* 105(8): 60–6.

95. Kang N, Tsang VT, Gallivan S, et al. 2006. Quality assurance in congenital heart surgery. Eur *J Cardiothorac Surg* 29(5): 693–7; discussion 97–8.

96. The Dartmouth Atlas Project. URL: http://www.dartmouthatlas.org/ (accessed September 2015). Dartmouth: The Dartmouth Institute for Health Policy and Clinical Practice.

97. Health Quality & Safety Commission. 2015. *NZ Atlas of Healthcare Variation.* URL: www.hqsc.govt.nz/our-programmes/health-quality-evaluation/projects/atlas-of-healthcare-variation.

98. Jackson G, Dalbeth N, Te Karu L, et al. 2014. Variation in gout care in Aotearoa New Zealand: a national analysis of quality markers. *NZ Med J* 127(1404): 37–47.

99. Dickie EE, Simcock JW. 2013. Reduction mammaplasty and resource allocation – are patients being treated fairly? An examination of the current New Zealand situation, and looking towards the future. *NZ Med J* 126(1374): 46–55.

100. Connolly MJ, Kenealy T, Barber PA, et al. 2011. National variability in provision of health services for major long-term conditions in New Zealand (a report from the ABCC NZ study). *NZ Med J* 124(1344): 16–35.

101. Wells S, Jackson R. 2011. Unwarranted variation in healthcare organisation and practice for long-term conditions. *NZ Med J* 124(1344): 6–9.

102. Raymont A. 2008. Hospital discharges in New Zealand 1991–2005: changes over time and variation between districts. *NZ Med J* 121(1279): 66–74.

103. Love T, Crampton P, Salmond C, et al. 2005. Patterns of medical practice variation: variability in referral for back pain by New Zealand general practitioners. *NZ Med J* 118(1212): U1381.

104. Harrison A. 2004. Provision of rheumatology services in New Zealand. *NZ Med J* 117(1192): U846.

105. Ellis C, Devlin G, Matsis P, et al. 2004. Acute Coronary Syndrome patients in New Zealand receive less invasive management when admitted to hospitals without invasive facilities. *NZ Med J* 117(1197): U954.

106. Paterson R. 2012. *The Good Doctor.* Auckland: Auckland University Press.

107. Shekelle PG, Pronovost PJ, Wachter RM, et al. 2013. The top patient safety strategies that can be encouraged for adoption now. *Ann Intern Med* 158(5 Pt 2): 365–8.

108. Tavare A. 2012. Where are we with transparency over performance of doctors and institutions? *BMJ* 345: e4464.

109. Hannan EL, Kilburn Jr. H, Racz M, et al. 1994. Improving the outcomes of coronary artery bypass surgery in New York State. *JAMA* 271(10): 761–6.

110. Lansky D. 1998. Measuring what matters to the public. *Health Aff* (Millwood) 17(4): 40–1.

111. Schwartz LM, Woloshin S, Birkmeyer JD. 2005. How do elderly patients decide where to go for major surgery? Telephone interview survey. *BMJ* 331(7520): 821.

112. Shahian DM, Normand SL. 2012. Autonomy, beneficence, justice, and the limits of provider profiling. *J Am Coll Cardiol* 59(25): 2383–6.

113. Clarke S, Oakley J. 2004. Informed consent and surgeons' performance. *J Med Philos* 29(1): 11–35.

114. Brown DL, Clarke S, Oakley J. 2012. Cardiac surgeon report cards, referral for cardiac surgery, and the ethical responsibilities of cardiologists. *J Am Coll Cardiol* 59(25): 2378–82.

115. NHS Choices. URL: http://www.nhs.uk/Pages/HomePage.aspx (accessed September 2015).

116. Khan OA, Iyengar S, Pontefract DE, et al. 2007. Impact of surgeon-specific data reporting on surgical training. *Ann R Coll Surg Engl* 89(8): 796–8.

117. Royal New Zealand College of General Practitioners. 2015. *Transparency of information related to medical/service interventions – College submission.* Wellington: Royal New Zealand College of General Practitioners.

118. Bardach NS, Hibbard JH, Greaves F, et al. 2015. Sources of traffic and visitors' preferences regarding online public reports of quality: web analytics and online survey results. *J Med Internet Res* 17(5): e102.

119. Emmert M, Meier F, Pisch F, et al. 2013. Physician choice making and characteristics associated with using physician-rating websites: cross-sectional study. *J Med Internet Res* 15(8): e187.

120. Grabner-Kräuter S, Waiguny MKJ. 2015. Insights into the impact of online physician reviews on patients' decision making: randomized experiment. *J Med Internet Res* 17(4): e93.

121. Mannion R, Davies HTO. 2002. Reporting health care performance: learning from the past, prospects for the future. *J Eval Clin Pract* 8(2): 215–28.

122. The Lancet. 2011. Public reporting of surgical outcomes. *The Lancet* 377(9772): 1126.

123. Davis MM, Hibbard JH, Milstein A. 2007. Consumer tolerance for inaccuracy in physician performance ratings: one size fits none. *Issue Brief Cent Stud Health Syst Change* (110): 1–5.

124. Friedberg MW, Damberg CL. 2012. A five-point checklist to help performance reports incentivize improvement and effectively guide patients. *Health Affairs* 31(3): 612–18.

125. Fung CH, Lim YW, Mattke S, et al. 2008. Systematic review: the evidence that publishing patient care performance data improves quality of care. *Ann Intern Med* 148(2): 111–23.

126. Accident Compensation Corporation. 12 August 2015. *Comments on HQSC Paper on Quality Reporting.* Wellington: Accident Compensation Corporation.

127. Casalino LP, Elster A, Eisenberg A, et al. 2007. Will pay-for-performance and quality reporting affect health care disparities? *Health Affairs* 26(3): w405–w14.

128. Brown DL, Clarke S, Oakley J. 2012. Cardiac surgeon report cards, referral for cardiac surgery, and the ethical responsibilities of cardiologists. *J Am Coll Cardiol* 59(25): 2378–82.

129. Brown DL, Epstein AM, Schneider EC. 2013. Influence of cardiac surgeon report cards on patient referral by cardiologists in New York State after 20 years of public reporting. *Circulation: Cardiovascular Quality and Outcomes* 6(6): 643–8.

130. Rothberg MB, Benjamin EM, Lindenauer PK. 2009. Public reporting of hospital quality: Recommendations to benefit patients and hospitals. *J Hosp Med* 4(9): 541–5.

131. Black N. 2013. Patient reported outcome measures could help transform healthcare. *BMJ* 346: f167.

132. Health Quality & Safety Commission. 2014. *Patient experience survey – methodology and procedures.* URL: www.hqsc.govt.nz/our-programmes/health-quality-evaluation/publications-and-resources/publication/1658.

133. Fanjiang G, von Glahn T, Chang H, et al. 2007. Providing patients web-based data to inform physician choice: if you build it, will they come? *J Gen Intern Med* 22(10): 1463–6.

134. Murphy PJ. 2012. Measuring and recording outcome. *Br J Anaesth* 109(1): 92–8.

135. Auras S, de Cruppé W, Blum K, et al. 2012. Mandatory quality reports in Germany from the hospitals' point of view: a cross-sectional observational study. *BMC Health Services Research* 12(1): 378.

136. Society for Cardiothoracic Surgery in Great Britain & Ireland (SCTS). National Adult Cardiac Surgical Database. URL: http://www.scts.org/ (accessed September 2015).

137. Fine LG. 2003. How to evaluate and improve the quality and credibility of an outcomes database: validation and feedback study on the UK Cardiac Surgery Experience. *BMJ* 326(7379): 25–8.

138. Chou E, Abboudi H, Shamim Khan M, et al. 2015. Should surgical outcomes be published? *J R Soc Med* 108(4): 127–35.

139. Epstein RM. 2002. Defining and assessing professional competence. *JAMA* 287(2): 226.

140. Dubinsky I, Jennings K, Greengarten M, et al. 2010. 360-degree physician performance assessment. *Healthc Q* 13(2): 71–6.

141. *Health Practitioners Competence Assurance Act*. 2003: New Zealand.

142. Exworthy M, Smith G, Gabel J, et al. 2010. Disclosing clinical performance: the case of cardiac surgery. *J Health Organ Manag* 24(6): 571–83.

143. Solberg LI, Mosser G, McDonald S. 1997. The three faces of performance measurement: improvement, accountability, and research. *Jt Comm J Qual Improv* 23(3): 135–47.

144. O'Leary DS. 1995. Measurement and accountability: taking careful aim. *Jt Comm J Qual Improv* 21(7): 354–7.

145. Chassin MR, Hannan EL, DeBuono BA. 1996. Benefits and hazards of reporting medical outcomes publicly. *N Engl J M* 334(6): 394–8.

146. Hafner JM, Williams SC, Koss RG, et al. 2011. The perceived impact of public reporting hospital performance data: interviews with hospital staff. *Int J Qual Health Care* 23(6): 697–704.

147. Tu JV, Donovan LR, Lee DS, et al. 2009. Effectiveness of public report cards for improving the quality of cardiac care: the EFFECT study: a randomized trial. *JAMA* 302(21): 2330–7.

148. Dahlke AR, Chung JW, Holl JL, et al. 2014. Evaluation of initial participation in public reporting of American College of Surgeons NSQIP surgical outcomes on Medicare's Hospital Compare website. *J Am Coll Surg* 218(3): 374–80.e5.

149. Glance LG, Dick A, Mukamel DB, et al. 2007. 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* 43(1p1): 300–12.

150. Jha AK, Epstein AM. 2006. The predictive accuracy of the New York State coronary artery bypass surgery report-card system. *Health Affairs* 25(3): 844–55.

151. Mukamel DB, Mushlin AI. 1998. Quality of care information makes a difference. *Medical Care* 36(7): 945–54.

152. Mukamel DB, Mushlin AI, Weimer D, et al. 2000. Do quality report cards play a role in HMOs' contracting practices? Evidence from New York State. *Health Serv Res* 35(1 Pt 2): 319–32.

153. Mukamel DB, Weimer DL, Zwanziger J, et al. 2004. Quality report cards, selection of cardiac surgeons, and racial disparities: a study of the publication of the New York State cardiac surgery reports. *Inquiry* 41(4): 435–46.

154. Mukamel DB, Weimer DL, Zwanziger J, et al. 2002. Quality of cardiac surgeons and managed care contracting practices. *Health Serv Res* 37(5): 1129–44.

155. Epstein AJ. 2010. Effects of report cards on referral patterns to cardiac surgeons. *J Health Econ* 29(5): 718–31.

156. Wang J, Hockenberry J, Chou S-Y, et al. 2011. Do bad report cards have consequences? Impacts of publicly reported provider quality information on the CABG market in Pennsylvania. *J Health Econ* 30(2): 392–407.

157. Sofaer S, Crofton C, Goldstein E, et al. 2005. What do consumers want to know about the quality of care in hospitals? *Health Serv Res* 40(6p2): 2018–36.

158. Maytham G, Kessaris N. 2011. A change in opinion on surgeon's performance indicators. *Interactive cardiovascular and thoracic surgery* 12(4): 586–90.

159. Hannan EL, Kilburn H, Lindsey ML, et al. 1992. Clinical versus administrative data bases for CABG surgery. *Med Care* 30(10): 892–907.

160. Hannan EL, Racz MJ, Jollis JG, et al. 1997. Using Medicare claims data to assess provider quality for CABG surgery: does it work well enough? *Health Serv Res* 31(6): 659–78.

161. Hannan EL. 1990. Adult open heart surgery in New York State. *JAMA* 264(21): 2768.

162. Dziuban SW, McIlduff JB, Miller SJ, et al. 1994. How a New York cardiac surgery program uses outcomes data. *Ann Thorac Surg* 58(6): 1871–6.

163. Dimick JB, Ghaferi AA, Osborne NH, et al. 2012. Reliability adjustment for reporting hospital outcomes with surgery. *Annals of Surgery* 255(4): 703–07.

164. California Cardiac Surgery and Intervention Project (CCSIP).
URL: http://www.californiacardiacsurgery.com/CCSIP2014/ (accessed September 2015).
California: California Society of Thoracic Surgeons (CASTS).

165. Ray S, Simpson I. 2012. Professional societies can lead the way on transparency but will need support. *BMJ* 345: e5075.

166. McCrum ML, Joynt K, Orav E, et al. 2013. Mortality for publicly reported conditions and overall hospital mortality rates. *JAMA Intern Med* 173(14): 1351.

167. Westaby S, De Silva R, Petrou M, et al. 2014. Surgeon-specific mortality data disguise wider failings in delivery of safe surgical services. *Eur J Cardiothorac Surg* 47(2): 341–5.

168. Aylin P, Alexandrescu R, Jen MH, et al. 2013. Day of week of procedure and 30 day mortality for elective surgery: retrospective analysis of hospital episode statistics. *BMJ* 346: f2424.

169. Armour Forse R, Bramble JD, McQuillan R. 2011. Team training can improve operating room performance. *Surgery* 150(4): 771–8.

170. Weaver SJ, Dy SM, Rosen MA. 2014. Team-training in healthcare: a narrative synthesis of the literature. *BMJ Qual Saf* 23(5): 359–72.

171. Merry AF, Weller J, Mitchell SJ. 2014. Improving the quality and safety of patient care in cardiac anesthesia. *J Cardiothorac Vasc Anesth* 28(5): 1341–51.

172. Wahr JA, Prager RL, Abernathy JH 3rd, et al. 2013. Patient safety in the cardiac operating room: human factors and teamwork: a scientific statement from the American Heart Association. *Circulation* 128(10): 1139-69.

173. Martinez EA, Shore A, Colantuoni E, et al. 2011. Cardiac surgery errors: results from the UK National Reporting and Learning System. *Int J Qual Health Care* 23(2): 151-8.

174. Mazzocco K, Petitti DB, Fong KT, et al. 2009. Surgical team behaviors and patient outcomes. *Am J Surg* 197(5): 678-85.

Appendix 1: Search strategy

As well as the formal literature search outlined below, sources were obtained using retrieval and review of referenced papers in major review articles. Further generation of retrieval and review of cited sources was undertaken from relevant papers found in the formal search.

Database: Ovid MEDLINE(R) In-Process & Other Non-Indexed Citations and Ovid MEDLINE(R) <1946 to Present> (adapted for Cochrane)

Search strategy:

1. ('public report*' or (public adj3 'performance data') or 'patient outcome report*').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] (996)
2. (Surgeon* or Individual* or Clinician* or Doctor* or Physician* or surgery or surgical).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] (3179575)
3. exp physicians/ or surgeons/ (90701)
4. 2 or 3 (3182839)
5. 1 and 4 (409)
6. ((disclos* or releas* or report* or provid*) adj3 ((physician* or doctor* or clinician* or doctor* or surgeon*) adj2 (data* 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] (399)
7. ((physician* or doctor* or clinician* or surgeon*) adj3 (rating* 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] (4799)
8. 1 and 7 (42)
9. Quality Improvement/ or 'Quality of Health Care'/ or Quality Assurance, Health Care/ or Quality Indicators, Health Care/ (122885)
10. Treatment Outcome/ or Patient Outcome Assessment/ or 'Outcome Assessment (Health Care)'/ or 'Outcome and Process Assessment (Health Care)'/ (751652)
11. Healthcare Disparities/ (8734)
12. Health Services Accessibility/ (54211)
13. 9 or 10 or 11 or 12 (913187)
14. 7 and 13 (1208)
15. 8 or 14 (1228)
16. 1 and 13 (558)
17. ('public reporting' or 'public report card*').m_titl. (300)
18. Consumer Behavior/ (17958)
19. Consumer Participation/ (14492)
20. ('public report*' or 'public data').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] (1587)

21. ((consumer or patient*) adj3 (choice or behaviour or behavior)).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] (37673)
22. 18 or 19 or 20 or 21 (53108)
23. 15 and 22 (73)
24. 5 or 6 or 8 or 16 or 17 or 23 (1205)
25. limit 24 to (english language and yr='1999 -Current') (1065)

Appendix 2: List of respondents that gave feedback

- Accident Compensation Corporation
- Association of Salaried Medical Specialists
- Australian and New Zealand College of Anaesthetists
- Consumer representatives
- Counties Manukau DHB Clinical Directors
- Ministry of Health
- National Chief Executives
- National Council of Medical Colleges group
- New Zealand Medical Association
- New Zealand Nurses Organisation
- New Zealand Public Service Association
- Royal Australasian College of Physicians
- Royal Australasian College of Surgeons
- Royal New Zealand College of General Practitioners
- Southern DHB clinicians
- Waitemata DHB

Appendix 3: Submission from Kevin Salmon

REPORT ON HEALTH & QUALITY COMMISSION AND MINISTRY OF HEALTH ONE DAY SEMINAR HELD IN WELLINGTON ON FRIDAY 3RD JULY 2015.

I arrived slightly late after flying in from Whangarei, having missed most of the introductions it was still clear that there were plenty of talented representatives from the medical profession but an equal number of consumer representatives from around the country.

I think it's important to note that I thought the discussion was around *the release of data on a surgeon's performance for those people considering elective surgery*. This has all arisen from a request by a reporter under the Official Information Act requesting how many operations a particular surgeon had performed.

After a quick recap of the two previous meeting held around the country (which I didn't attend) there was a very informative talk from Andrew Connelly – Chairman of the NZ Medical Council, also Richard Hamblin spoke about data limitations and Karen Evison (Ministry of Health) spoke about the cardiac register which costs around $2m a year to run. It also proved statistically that the best place in the country to have a heart attack was in Dunedin. I thought to myself, that's sad I live in Whangarei.

We spend the day in groups discussing *"what does the consumer want, in terms of surgery"*. Amongst so many capable people it was, at times hard to get an opinion in but I think I got there in the end.

Overview

We spent so much time talking about how data is gathered, how limited is and how it could be used to improve services BUT did not focus on how the poor consumer would be able to access this information, how they would understand it and in fact if all of that time and effort really added any value to the consumer at all.

Important points:

- Can you provide information on a specific surgeon about his competency to do an operation when so many other people are involved? Especially around multiple disciplinary teams making treatment plans for patients, when it goes wrong is it one person's "fault". Is the anaesthetist subject to the same scrutiny as a surgeon?
- The data even when accurately gathered can be manipulated so much it can be made to look as good as you wish.
- Will the release of such information have a negative effect on surgeons performing difficult surgeries, will they just not attempt them if makes them look like they have poor results?
- What is a poor result, staying alive vs quality of life? If somebody dies at home 30 days after surgery is that the responsibility of the surgeon?
- Is this all so complicated it really is going to be almost impossible to provide useful information to a consumer that they can make an informed decision?

Case study

My brother, Bob had been on a waiting list for a replacement hip for just on two years, he was in great pain and bordering on being unable to work as a builder, he was advised he was now on a the short list for surgery and it would happen within the next six months. True to that information he was given a date for the operation (four days after our Mum had passed away), the operation was a huge success, he is back at work, playing golf again and enjoying free movement.

My point is the assumption is we have “a choice” in reality we do not. Imagine if he had accessed his surgeon’s record of hip replacements, seen it was rated poorly by consumers. Would he have said “no” I want somebody else please? He would have had his place taken by the next person on the list and he would now be in constant pain and on a sickness benefit. Or another scenario, having seen the best hip surgeon in the country is in Christchurch and he would like him to do the operation. That truly has to be a joke.

What could happen

- Most of us trust the Medical Council to make sure surgeons are not only competent but able to perform an operation to the highest standard.
- The use of peer reviews, independent audits and annual appraisals would be included in that process to provide transparent records of competency.
- If a surgeon is under investigation or has been suspended that should be made available to the public via the Medical Council website.
- All information on a surgeon’s performance is available to the Council so extra training, supervision or restrictions could be applied as necessary. It is assumed the College of Surgeons doesn’t already administer this role
- The public needs to have confidence the surgeon is being monitored correctly and when something goes wrong the appropriate actions are taken quickly.

Summary

I am sure there is a percentage of the population that would wade through a vast amount of information, that has cost millions to put together (I assume at the cost of services not being provided) and satisfy themselves that that information is accurate, trustworthy and of value to them to make an informed decision about the surgeon who will be operating on them.

I believe the reality is with little or no choice in the matter, the consumer must trust “the system” to monitor all aspects of a surgeons competency and that the information that is gained be used by all DHB’s to assist with best practice models throughout the country so having a heart attack in Whangarei will give me the same opportunities to survive as having one in Dunedin.

Question

I wonder how Air New Zealand would react if the public started asking questions about “Is the pilot who was flying my plane today competent?”

Kevin Salmon *– Chairman, Northland DHB Consumer Council*

newzealand.govt.nz