This is the Author’s version of the paper published as:

Title: Surgeons' Report Cards, Heuristics, Biases and Informed Decision Making
Author: S. Clarke
Author Address: Stclarke@csu.edu.au
Year: 2007
Chapter: 11
Pages: 167-189
Book Title: Informed Consent and Clinician Accountability: The Ethics of Report Cards on Surgeon Performance
Editor: Publisher: Cambridge University Press
City: Cambridge
ISBN: 9780521865074
Publisher URL: http://www.cambridge.org/catalogue/catalogue.asp?isbn=9780521687782
Informed Consent and Clinician Accountability

The Ethics of Report Cards on Surgeon Performance

Edited by

Steve Clarke
University of Oxford, UK, and Charles Sturt University, New South Wales, Australia

Justin Oakley
Monash University, Victoria, Australia
Surgeons’ report cards, heuristics, biases and informed consent

Steve Clarke

1. Informed consent and surgeons’ performance information

An important reason for providing patients with performance data on individual surgeons is to enable patients to make better decisions about surgery, as a part of the informed consent process. Surgeons' performance data can be utilised to enable a variety of types of decision that a patient may face. A patient can utilise performance data on individual surgeons to enable a choice between available surgeons. A patient can utilise surgeons’ performance data when deciding between surgery involving an available surgeon and a non-surgical alternative form of treatment. Also, a patient can utilise surgeons’ performance data to help decide whether or not to wait for a high-performing surgeon, who is not currently available, to become available.

Traditionally, performance data on individual surgeons have not been disclosed to patients, and such data have not usually been thought necessary to disclose for the purposes of providing effective informed consent. Canonical treatments of the doctrine of informed consent, such as Faden and Beauchamp (1986), do not consider the possibility of making such information available to patients. However, it has recently been argued that the doctrine of informed consent implicitly requires that surgeons’ performance data be made available to patients (Clarke and Oakley, 2004). The gist of this argument is easy enough to grasp: It is uncontroversial that the
significant and material risks associated with an operation should be disclosed to a patient who is contemplating that operation. If an operation is known to involve a 10% risk of mortality and there is a failure to disclose that it involves a 10% risk of mortality, before the operation takes place, then effective informed consent has not been provided. However, the actual risks of an operation will vary, according *inter alia* to the performance ability of the surgeon conducting the operation. An operation that has a 10% risk of mortality, on average, may only have a 5% risk of mortality when conducted by one surgeon and a 15% risk of mortality when conducted by another surgeon. So a disclosure of the actual risks of an operation, for the purposes of enabling effective informed consent, needs to include information about the performance abilities of available surgeons.

In the United States and the United Kingdom, the two countries in which cardiac surgeons' performance data have been made publicly available, considerations other than a concern to enable effective informed consent have led to the publicising of individual surgeons' performance data. The motive for making cardiac surgeons’ performance data publicly available in the United Kingdom has been a desire to change the closed culture of medicine that led to the Bristol Royal Infirmary Scandal, and to ensure that high standards of professional accountability are now met (Neil et al., 2004, p. 266). In the United States an explicit aim of public release, which has taken place in New York, Pennsylvania and New Jersey, has been to improve clinical quality. This is to be done by providing information that can improve the quality of patients' decisions and then allowing market mechanisms to influence the behaviour of surgeons. (Neil et al., 2004, p. 267; Marshall et al., 2000).
In New York State, tables of surgeons’ comparative performance information (known colloquially as ‘report cards’) for coronary artery bypass graft (CABG) operations, the most common form of cardiac surgery, have been made publicly available for over fifteen years. If we look at the most recent report card on CABG (2001-2003) in New York State we find information such as the following:

**St. Francis Hospital:**

<table>
<thead>
<tr>
<th>Name</th>
<th>No. of Cases</th>
<th>No. of Deaths</th>
<th>OMR</th>
<th>EMR</th>
<th>RAMR</th>
<th>95% CI for RAMR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bercow N</td>
<td>697</td>
<td>27</td>
<td>3.87</td>
<td>2.35</td>
<td>3.36</td>
<td>(2.21, 4.89)</td>
</tr>
<tr>
<td>Colangelo R</td>
<td>751</td>
<td>17</td>
<td>2.26</td>
<td>2.33</td>
<td>1.97</td>
<td>(1.15, 3.16)</td>
</tr>
<tr>
<td>Damus P</td>
<td>501</td>
<td>5</td>
<td>1.00</td>
<td>1.73</td>
<td>1.17</td>
<td>(0.38, 2.74)</td>
</tr>
</tbody>
</table>


**OMR:** The observed mortality rate is the number of observed deaths within 12 months, of isolated CABG surgery, divided by the number of patients.

**EMR:** The expected mortality rate is the sum of the predicted probabilities of death for each patient, divided by the total number of patients.

**RAMR:** The risk-adjusted mortality rate is the best estimate of what the provider’s mortality rate would have been if the provider had a mix of patients identical to the statewide mix.

**95% CI for RAMR:** The 95% Confidence Interval (CI) for RAMR is indicative of the degree of confidence that we are warranted in attaching to RAMR figures for individual surgeons. N. Berkow has
a RAMR of 3.36 with a 95% CI (2.21, 4.89). This means that we can be 95% confident that his actual RAMR falls between 2.21 and 4.89.

Doctors operate at different hospitals, which receive different mixes of patients. And not all surgeons at a particular hospital will treat a similar mix of patients. More experienced surgeons, and surgeons whose ability is regarded as superior by their peers, may be asked to conduct more of the difficult, higher-risk operations that need to be conducted. To achieve a fair reflection of a surgeon’s performance, given his or her OMR, we need to adjust for the EMR of his or her particular mix of patients, producing the surgeon’s RAMR. Risk adjustments are made by considering a variety of factors that affect outcomes, such as age, gender, ventricular function and the presence of significant ‘comorbidities’.

In the United Kingdom risk-adjusted performance information for individual cardiac surgeons is also collated and made available to the public. These data were released in a very coarse-grained form, from 2004 to 2006, with British cardiac surgeons being rated with the use of a three point scale (Neil et al., 2004). However, Britain has now moved to a system of providing percentile risk adjusted survival rates (RASR) for individual cardiac surgeons conducting CABG operations and aortic valve replacement operations.¹ These data are not currently complete, with surgeons from only 17 of the 33 heart units in England and Wales now providing RASR for their operations. Figures will be updated annually and, although participation is voluntary, it is expected that these data will soon become much more comprehensive (Healthcare Commission 2006).
Here I consider evidence from the field of behavioural decision making, which suggests that ordinary decision making is affected by a variety of systematic biases, and I investigate some ways in which this research bears on issues concerning the presentation of comparative surgeons’ performance data for the purposes of enabling informed consent. These are biases that are particularly relevant to the interpretation of statistical information and which can be extremely serious. Consider a recent study by Yamagishi (1997). Most participants in this study of lay estimates of risk rated a cancer as riskier, when it was described as one that 'kills 1,286 out of 10,000 people’, than when it was described as one that ‘kills 24.14 out of 100 people’. But, in actual fact, a risk of 1,286 in 10,000 is approximately half as severe as a risk of 24.14 in 100. This result is an instance of the bias of base-rate neglect (Tversky and Kahneman, 1974). 1,286 is a large number relative to 24.14 and it appears that participants in this study directly compared these two numerators, while losing sight of the fact that they are intended to be understood in relation to different denominators.

Along with Faden and Beauchamp (1986, p. 235) and most other contributors to the literature, I hold that the overriding goal of the informed consent process is to uphold the value of patient autonomy. While the autonomous choices of patients need not be based solely on rational factors, a patient is not able to make a fully autonomous choice to provide consent, unless it is possible for that patient to rationally deliberate about the alternative courses of action available to her. But if psychological biases seriously erode a patient’s deliberative capacities, then that patient is not able to deliberate rationally and hence not fully able to act autonomously; and the key value that the doctrine of informed consent is designed to uphold can not be upheld. So it is crucial, if we are to enable effective informed consent, that we are sensitive to the
ways in which surgeons’ performance information is processed by patients and the
ways in which that processing may become distorted.

2. Heuristics, biases and dual-processing

The phenomenon of base-rate neglect can be explained by appeal to the activation of
the 'representativeness' heuristic (Tversky and Kahneman, 1974). Rather than
adjusting the two fractions to be compared to a common denominator, and then
making a comparison, most of Yamagishi's (1997) subjects looked for an implicit cue
indicative of the relative size of the two fractions. We can often make comparisons of
the relative size of numbers intuitively, just by looking at the number of digits
contained in those numbers. In one case we have a numerator with four digits, while
in the second case we have a numerator with only two digits (ignoring the digits after
the decimal point). Generally, numbers with more digits are larger numbers and we
may be inclined to apply this intuitive cue and judge the fraction containing the larger
of the two numerators to be the larger of the two fractions, instead of performing the
laborious calculations that would lead us to be able to directly compare the two.
Unfortunately, in some circumstances, the application of such quick and easy
heuristics produces the wrong result.

Another systematic bias that can be explained by the activation of the
representativeness heuristic is the ‘conjunction fallacy’, memorably described in
Tversky and Kahneman's (1983) 'Linda' example. In their (1983) study subjects were
given the following information:
Linda is 31 years old, single, outspoken and very bright. She majored in philosophy. As a student, she was deeply concerned with issues of discrimination and social justice, and also participated in anti-nuclear demonstrations. (Tversky and Kahnemann, 1983, p. 297)

Having read the above statement, subjects were asked to rank the probability of a set of eight descriptions, which included the following (along with six 'fillers'):

6. Linda is a bank teller.
8. Linda is a bank teller and is active in the feminist movement.

An application of elementary logic would lead to the conclusion that description 8 cannot be more probable than description 6. Description 8 is a conjunction, description 6 is one of its conjuncts, and a conjunction can never be more probable than one of its conjuncts. Nevertheless, 85% of the research subjects concluded that description 8 was more probable than description 6. Rather than reasoning logically, it appears that they approached the problem intuitively, applying implicit stereotypes. Linda represents the common stereotype of a ‘feminist’ much more closely than she represents the common stereotype of a ‘bank teller’, and so the representativeness heuristic was activated in 85% of the research subjects, who intuited that description 8 was more probable than description 6.

As well as various biases that have been explained by appeal to the representativeness heuristic, there are many other cognitive biases that psychologists have assembled evidence for and have been explained by appeal to the operation of other heuristics. These include biases explained by appeal to the 'availability heuristic' (Tversky and
Kahneman, 1974), in which the easy availability of certain sorts of information leads to people giving disproportionate weighting to that information, at the expense of less available information. There are also biases that have been explained by appeal to the 'affect heuristic', in which intuitive judgements of goodness or badness influence assessments of likelihood and influence a variety of other assessments (Slovic et al., 2002).

The conjunction fallacy is often committed by people who are capable of reasoning logically about conjunction, but who nevertheless apply heuristics, such as the representativeness heuristic. Why do we employ heuristics in our thinking, when these can lead us astray and when we could often avoid being led astray by applying systematic logical reasoning? The best answer to this question, that I am aware of, starts with the plausible assumption that our minds are ‘dual processors’. We employ two very different sorts of cognitive processes, a controlled effortful reasoning process and an automatic effortless intuitive process (Kahneman and Frederick, 2002). Dual-processing theorists Stanovich and West (2002) refer to automatic intuitive processes as instances of the application of 'System 1' and effortful reasoning processes as instances of the application of 'System 2'. It is probably not feasible to use System 2 processing much more than we do, in the view of most dual-processing theorists, because conscious attention, required for System 2 processing, but not for System 1 processing, is a precious resource, which we need to deploy sparingly (Moskowitz et al., 1999, pp. 26-30).

The operation of particular instances of System 1 and System 2 processing occur independently of one another. However, System 1 is an adaptive system, and complex
cognitive operations that form a part of System 2 may influence the development of
System 1 over the course of time. For example, a dedicated chess player can, over
time, acquire an ability to intuitively 'read' a position, using intuitions that have been
'trained up' by the repeated application of System 2 to the game of chess. Switching
between System 1 and System 2 processing generally takes place automatically and
often goes unrecognised. Even when we do recognise that we have been applying
System 1 processing, in circumstances where it may be more appropriate to employ
System 2 processing, and we then go on to apply System 2 processing, it may be hard
to prevent ourselves from automatically reverting to System 1 processing. Steven Jay
Gould’s account of the Linda case study as one in which ‘… a little homunculus in
my head continues to jump up and down, shouting at me — but she can't just be a bank
teller; read the description.’ (1991, p. 469), appears to be an example of such
automatic reversion taking place.

It is sometimes assumed that Tversky and Kahneman (1974; 1983) have demonstrated
that human judgements are frequently in error. But although it is easy enough to
demonstrate that some System 1 judgments are liable to systematic bias, there is a
long way to go if we want to establish that the use of System 1 processing leads us to
commit frequent errors of judgement in natural settings. Gerd Gigerenzer and his
colleagues have mounted a spirited defence of reasoning based on System 1
processing, against the charge of frequent error (Gigerenzer and Todd, 1999). They
argue, roughly, that heuristics are adaptive tools that enable us to employ effective
System 1 processing in situations that require rapid decision making, and in which it
would be impractical to employ System 2 processing to make decisions in a time-
effective manner.\(^5\) They also hold that the dramatic demonstrations of failures of
natural decision making, in the sorts of experiments that Tversky and Kahneman (1974; 1983), have made famous, are largely confined to the artificial experimental settings that they are conducted in.\(^6\)

The presentation of comparative surgeons’ performance information is one setting that appears to approximate closely to the artificial experimental settings that Tversky and Kahneman (1974; 1983) and others have examined. Comparative surgeons’ performance information is typically presented in an abstract mathematical format, as exemplified by the earlier extract from the New York State Department of Health (2005). So it seems very doubtful that Gigerenzer and Todd’s (1999) defence of the use of System 1 processing in natural settings – even if it were judged to be successful for typical natural settings – would be sufficient to erase our concerns about the use of System 1 reasoning in an unusual natural setting that approximates closely to the artificial experimental settings that Tversky and Kahneman’s experiments have mostly been conducted in. It is, of course, possible that in the future Gigerenzer and his colleagues, or some other researchers, will produce evidence that will be sufficient to enable us to set aside concerns about biased System I processing in the interpretation of comparative surgeons’ performance information. However, in the absence of such evidence, it seems prudent to err on the side of caution and assume that the lay interpretation of comparative surgeons’ performance information will be subject to systematic bias.

3. Responding to bias
A number of different policy responses to the problem of cognitive bias in patient’s interpretation of surgeons' performance information in the informed consent process have been suggested to me. One response is to seek to avoid bias by encouraging patients to employ System 2 processing when incorporating surgeons' performance information into their decision making processes for the purposes of informed consent. A second suggestion is to accept that patient’s interpretations of surgeons’ performance data will be biased, and encourage patients to rely on expert testimony about comparative surgeons' performance ability, when deciding whether or not to consent to be operated on by a particular surgeon.

A third response is to accept that patients will employ System 1 processing, when incorporating surgeons' performance information into their decisions to consent and to seek to avoid bias by translating the relevant statistical information into ordinary language, before presenting it to patients. A fourth option is to accept that patients will employ System 1 processing, when incorporating surgeons' performance information into their decisions to consent and to try to present that statistical information in such a way as to minimise the potential for error. In this section I consider these options in turn. I will argue that the first three are impractical and that option four is the most feasible course of action.

The first option is currently impractical because the level of statistical education amongst patients is low (Lloyd, 2001). Many will not be capable of applying System 2 reasoning to the problem of calculating how the performance ability of a particular surgeon modifies the level of risk associated with a particular operation. It would take a major overhaul of education systems in most countries to improve the statistical
education of the general public to the level where most are capable of performing such calculations. While improving statistical education is a worthy aim, it is not a viable solution to our problem in the short or even medium term. But even if we could improve statistical education to the desired level, it would still not ensure that most people would employ System 2 processing to incorporate surgeons’ performance information into their decisions to provide informed consent for surgery. Even people who are statistically educated tend to employ System 1 processing to ‘extract the gist’ of risk information, rather than utilise System 2 processing, when attempting to understand risk information (Lloyd et al., 2001).

It can be surprisingly difficult to convince people that their use of reasoning based on System 1 processing has biased their decisions. Although people are able to recognise that others are frequently the victims of cognitive bias, they are often oblivious to the possibility that they themselves may be victims of bias. Furthermore, there is no guarantee, even if they are convinced, that they will not slide back into System 1 modes of reasoning as Steven Jay Gould repeatedly did, when faced with the Linda example. Wilson and Brekke (1994, pp. 119-120) argue that for a person to successfully avoid biased reasoning four conditions must be met. First, they must be aware of the presence of a biased mental process. Second, they must be motivated to correct any error that the biased mental process has caused. Third, they must be aware of the direction and the magnitude of the error. An over-correction or under-correction of an error results only in further error. Fourth, even if the first three conditions are satisfied, they must be able to exert sufficient control over their mental processes to correct the error. Wilson and Brekke (1994) argue that it is exceedingly difficult for all four of these conditions to be met.
Asking patients to rely on expert testimony, our second option, would be vigorously resisted by orthodox interpreters of the doctrine of informed consent, such as Faden and Beauchamp (1986), who emphasise the importance of comprehensive disclosures in the informed consent process. I am not opposed to the utilisation of expert testimony, as a part of the informed consent process, for reasons that are set out in Clarke (2001). However, the appeal to expert testimony as a substitute for comprehension may be impractical to implement in this context. Doctors are the standard providers of disclosure of relevant information. Doctors are generally somewhat better educated about the interpretation of statistics than lay folk and can play a part in helping patients to understand basic statistics (Paling, 2003). Nevertheless, like patients, doctors are often in the grip of cognitive biases, in virtue of the fact that they typically apply System 1 processing to the interpretation of statistical information (Schwartz, 1994). So, relying on doctors’ testimony would not solve the problem of System 1 processing-induced bias in the interpretation of statistics. It might be possible to involve trained statisticians and experts in behavioural decision making in the informed consent process, but given how few of these there currently are, who could reasonably be expected to be available, and give the expenses involved in employing such experts this seems a very utopian solution to our general problem, so it is one that will not be investigated here.

The third option, translation into ordinary language, also turns out to be impractical because of the vagueness of ordinary language. When people are asked to quantify ordinary probabilistic terminology, they typically provide a very broad range of answers (Mosteller and Youtz, 1990). This is not only true of lay interpreters.
Mosteller and Youtz (1990, p. 3) discuss a case in which four experts were asked to quantify the phrase 'a very real possibility'. Their respective answers were 2%, 10%, 35% and 'less than even'. But even if conventions could be established about the translation of statistics into lay terminology, there is not much reason to believe that the biases that plague the interpretation of statistics would thereby be avoided. The failure of lay interpreters to understand the effect of conjunction, in the 'Linda' study, was not a consequence of the information being presented in a statistical format. Rather, it was a consequence of their failure to employ elementary logical considerations, in circumstances where it seems that elementary logical considerations should have been employed. Shifting into ordinary language format does not obviate the need for the employment of logical considerations when reasoning about the Linda case, but there is no reason to suppose that such a shift would prompt people to employ ordinary logical considerations in their deliberations. The biases in question are, it seems, a by-product of System 1 processing, which is highly likely to be operative in manipulations of information, regardless of how that information is presented. So it seems that there is little reason to hope that the use of ordinary language can provide a way to avoid cognitive biases, in the processing of information about risk.

Given the propensity of people to employ System 1 processing to interpret statistics and make decisions based on statistics, it seems best to investigate our fourth option and accept that patients will typically employ System 1 processing when interpreting surgeons' performance information and incorporating this information into their decision making processes. For the purposes of obtaining informed consent, we should ensure that surgeons' performance data are presented in such a way as to
enable people to make the best decisions that they can, using System 1 processing. For the benefit of those patients who are willing and able to make consenting decisions incorporating surgeons' performance information, on the basis of System 2 processing, we should make surgeons' performance information available in a second format, suitable for System 2 processing. However, the initial presentation of information, which forms a part of the informed consent process, should be tailored to System 1 processing.

4. Presentation of data

For the purposes of enabling accurate System 1 processing, we should present individual surgeons’ performance information to patients in a simple and clear format that does not prompt the application of unnecessary System 1 heuristics, which may lead to error. The OMR and the EMR, presented in the New York State report cards, are unnecessary for the purposes of patient decision making and may prompt erroneous inferences. Patients’ tendencies to ‘extract the gist’ of risk information (Lloyd et al., 2001), means that they are liable to make decisions on the basis of a mix of OMR, EMR and RAMR (or RASR), instead of focussing solely on RAMR (or RASR). So we should only present RAMR, or RASR, as is now done in the United Kingdom. Of course there is no objection to including OMR and the EMR in a System 2 format report card.

Explicit rankings of individual surgeons, such as those that occur in the New York State report cards, are also unhelpful to patients in circumstances where there is
considerable overlap in the 95% confidence intervals of those individuals (as there is in our earlier excerpt from the New York State report cards). Many patients will not understand the concept of a 95% confidence interval and may be liable to draw the erroneous conclusion that, if one surgeon has a lower RAMR than a second, then he or she is definitely a superior surgeon to the second surgeon, even if there is considerable overlap between the 95% confidence intervals for their respective RAMRs. A better alternative for patients is an easy-to-comprehend broad banding system of ranking individuals.¹¹

The three point scale, which was used to rank the performance ability of cardiac surgeons in the United Kingdom from 2004 to 2006, before percentile performance data for individual surgeons began to be provided in April 2006, was an easy-to-comprehend broad banding system of ranking individuals. However, it was probably too coarse-grained for the purposes of enabling surgeons’ performance information to be fully incorporated into the informed consent process. If we are warranted in being confident that, for a given type of operation, one surgeon is a better performing surgeon than a second, then it is important that that information be presented to patients, for the purposes of enabling effective informed consent. It may well be that, within a given performance band, there were some British cardiac surgeons whose performance ability was demonstrably superior to other British cardiac surgeons, who were listed as being within the same performance band. If some surgeons within the same band are demonstrably superior performers to other surgeons, within the same band, then it seems that we should provide finer-grained information to patients for the purposes of better enabling effective informed consent. In such circumstances, the provision of finer-grained information would not introduce any new interpretive
biases, or increase the magnitude of existing biases. However, it will enable patient to
decide more effectively when choosing between surgeons, and when choosing
between surgery and non-surgical alternatives.\textsuperscript{12}

References


Clarke, S. (2001). Informed consent in medicine in comparison with consent in other


Gigerenzer, G. and Todd, P. M. (1999). Fast and frugal heuristics: the adaptive
toolbox. In *Simple Heuristics that Make Us Smart*, ed G. Gigerenzer, P. M. Todd and
the ABC Research Group. New York: Oxford University Press, pp. 3-34.


http://www.healthcarecommission.org.uk/newsandevents/pressreleases.cfm?cit_id=3785&FAArea1=customWidgets.content_view_1&usecache=false


**Endnotes**

1 The choice of presenting surgeons’ performance information in terms of survival rather than mortality rates can be expected to have a significant effect on the behaviour of prospective patients, in virtue of the influence of ‘framing effects’. In general, we can expect that the presentation of such data in terms of mortality rates...
will encourage more risk-averse behaviour in prospective patients than will presentation in terms of survival rates (See Kahneman and Tversky, 1984). Although different frames will promote different values, in general there is no one correct way to frame data, any more than there is a correct way to frame a picture. Framing does not typically involve a distortion from accuracy and there are not generally any correct frames. So, framing is conceptually distinct from bias. Discussion of the various philosophical issues raised by framing is beyond the scope of this chapter.

2 I am not alone in arguing for the importance of consideration of psychological biases, in the articulation of the informed consent process. See also Thompson (1996) and Lloyd et al., (2001).

3 I have focussed specifically on issues of presentation of comparative surgeons’ performance data to patients. There is a growing body of literature that addresses more general issues of information presentation, for the purpose of informing patient choice. See for example Hibbard et al. (1997); Hibbard (2003).

4 For more on the relations between autonomy, informed consent and deliberation see Beauchamp and Childress (2001, pp. 57-98).

5 Some of the heuristics that Gigerenzer et al. (2002) discuss may not be instances of System 1 processing, but consciously applied 'rules of thumb', which look more like special cases of System 2 reasoning.
6 For a discussion of these and other lines of criticisms of the ‘heuristics and biases’ research program, see Gilovich and Griffin (2002).

7 For a discussion of when it is appropriate to employ System 1 processing, and when it is appropriate to employ System 2 processing, see Kleinmuntz (1990).

8 I do not wish to deny that our second option could be used on an ad hoc basis. My claim is it is impractical to apply it systematically and it will remain impractical to do so for the foreseeable future.

9 It is sometimes suggested that the presentation of risk information in a frequency format can eliminate bias. However, see Gilovich and Griffin (2002, pp. 14-15).

10 Thanks to Justin Oakley for this suggestion.

11 Note that a broad banding system still requires representation of a measure of uncertainty (Royal Statistical Society Working Party on Performance Monitoring in the Public Services, 2005).

12 Thanks to Justin Oakley, Steve Matthews and audiences at the Centre for Applied Philosophy and Public Ethics, Canberra and Wagga Wagga divisions, and at 'Publicising Performance Data on Individual Surgeons: the Ethical Issues', a workshop sponsored by the Academy of the Social Sciences in Australia, held in Melbourne 2004. This research was supported by National Health and Medical Research Council Project Grant 236877.