The Healthcare Commission has inherited the ratings task for 2003-4 and has recently advised that the number of breaches will now be measured as a percentage of the total number of patients a trust treats.² Our analysis shows that trusts need to monitor constantly the statistical rigour of the performance indicators used to assess them.

We thank L Parker for help with the statistical analyses and M F Laker for advice.

Contributors and sources: RMB is an executive director with extensive knowledge and experience of planning and performance management in healthcare. MSP is a Royal Statistical Society recognised chartered statistician and an experienced non-clinical epidemiologist. MI is a retired clinical academic with a background in health technology assessment and the transfer of evidence based practice into acute trust policies. All the above contributed equally to the paper. The article resulted from a detailed analysis of the reasons why a high performing trust breached a key indicator in the 2002-3 star ratings exercise.

Competing interests: RMB and MI are members of the Newcastle upon Tyne Hospitals NHS Trust Board of Directors. MSP was paid by the trust for a statistical report on the trust's performance in the star ratings exercise

Summary points

Performance of NHS trusts is reported in terms of a star rating

The methods used to assess performance in 2002-3 were flawed

Use of absolute rather than relative figures for some measures created bias against larger trusts

Use of end of quarter figures for outpatient waiting times masked large breaches in some hospitals

More statistical rigour is needed to ensure trusts' confidence in the ratings system

1 Bird SM, Cox D, Farewell VT, Goldstein H, Holt T, Smith PC. *Performance indicators: good, bad, and ugly*. London: Royal Statistical Society, 2003.

2 Healthcare Commission. *Indicator listings for acute trusts*. http://ratings.healthcarecommission.org.uk/Indicators_2004/ (accessed 11 July 2004).

(Accepted 18 April 2004)

Commentary: Performance assessment is here to stay

Anna Walker

Barker and colleagues look at how performance on one indicator out of 45 affected last year's star rating for their trust. The two specific concerns they raise (which the Commission for Healthcare Improvement discussed with them last summer) are that the target was measured at the quarter end and that it was measured using absolute rather than relative thresholds.

No one has ever argued that the commission's method was the only way to measure the target. When the commission accepted responsibility for ratings from the Department of Health, some consistency had to be maintained. The previous year, the Department of Health had assessed performance on both inpatient and outpatient targets on an absolute basis.

The absolute thresholds were retained because the target is framed in absolute terms—that is, no patient will wait longer than a specified time. Quarter end data were chosen so that they could be adjusted to exclude, for example, Welsh patients treated in English trusts, who are not subject to the same rigorous waiting times targets. The data used were the only set for which such an adjustment was made by the Department of Health in 2002-3.

Trusts had an opportunity to comment on the methods used to construct all the indicators included in ratings before they were finally published. Two or three trusts commented on the construction of this indicator, but Newcastle did not make these points until after the ratings were published.

The commission committed itself to a review after it had published its first set of ratings in 2003. We now know the result of that review, and the outpatient target will be measured this year by using month end data and with proportionate thresholds, meeting many of the concerns that the Newcastle case illustrates. This revised approach, however, would be unlikely to have

changed the result for Newcastle last year. According to its own quarter end data, calculated in proportion to overall patient numbers, Newcastle had the sixth high percentage of breaches in the country.

Future of performance ratings

Although there is always benefit in looking at the past, I am keen to look to the future on assessing the performance of NHS organisations. Everyone has always agreed that star ratings are not perfect but also that performance assessment can be beneficial. It is now time to build some agreement about a better method.

The Healthcare Commission wants a system of performance assessment that is more accessible to the public, that drives improvement in the NHS, that is seen as relevant and fair by the service and clinicians, and that is more comprehensive in how it measures organisations. If our performance assessments are going to be used to decide rewards and sanctions, we also need to work with the Department of Health to ensure that they are fit for that purpose.

This year's and next year's star ratings will look familiar. We have made and will make improvements, but some of the aspects that have been criticised will necessarily remain. As a new organisation we are beginning to consider how we will assess NHS organisations in the future. In the autumn, we will publish a consultative document about our future methods of assessment, with a view to using the new methods for ratings from 2006. We will be looking to engage with doctors, nurses, and others across the NHS about how this system might work. I hope *BMJ* readers will take the opportunity to help us design something better.

Competing interests: None declared.

Healthcare Commission,
London EC1Y 8TG
Anna Walker
chief executive
anna.walker@healthcarecommission.org.uk