

FEATURE

MID STAFFS INQUIRY

How the message from mortality figures was missed at Mid Staffs

nigel.hawkes1@btinternet.com

Do published outcomes tell us what's really going on inside a hospital? **Nigel Hawkes** reports on how standardised mortality ratios at Mid Staffordshire NHS Foundation Trust came to conceal the hospital's failings

Nigel Hawkes *freelance journalist*

London, UK

The new NHS in England aims to empower the patient and shame failing hospitals by measuring and publishing outcomes. Open data will make health providers honest and patients better able to choose between them. How simple is that?

Rather too simple, to judge by evidence presented to the public inquiry into Mid Staffordshire NHS Foundation Trust, which has shown that the old NHS had ingenious ways of burnishing its outcomes and concealing its failings. Even death, the most unambiguous outcome of all, lost its sting when changes in coding practice resulted in many deaths being excluded from the calculation of hospital mortality ratios.

Whether this amounted to actual dishonesty may never be determined. In his final submission Tom Kark, counsel to the inquiry, suggested that it may not be necessary for Robert Francis, the inquiry's chair, to make any finding of fact over claims that Mid-Staffs and two other West Midlands trusts engaged in a plan to lower their apparent mortality ratios by changing the way patients were coded.¹ Nor did Kark reach any firm conclusion over the suggestion that West Midlands Strategic Health Authority (SHA) sought to muddy the waters by commissioning a critical report into mortality measures, published in the *BMJ*, from academics known to disbelieve in the method.

Both these charges were made in evidence to the inquiry by Professor Brian Jarman, the father of the hospital standardised mortality ratio (HSMR, see box), which has been used for the past decade by the health analytics company Dr Foster to compare hospital performance. By his reckoning, West Midlands tried to undermine the mortality data in two ways: firstly, by discrediting the method and, secondly, by encouraging three hospital trusts in its area (Mid Staffs, Walsall, and George Eliot) to improve their ratings by classifying increasing numbers of patients as palliative care cases. Since hospitals cannot be accused of poor care if they do not save the lives of people admitted to die, the HSMR is adjusted for patients coded as needing palliative care. The more patients so coded, the lower (better) the HSMR.

Credibility of data

Item one on the charge sheet is a report, *Probing variations in hospital standardised mortality ratios in the West Midlands*, edited by Mohammed Mohammed and Richard Lilford of the Department of Public Health and Epidemiology at Birmingham University. It was commissioned by West Midlands SHA, and its steering committee was chaired by Rashmi Shukla, medical director of the SHA, and heavily weighted with representatives from hospitals in the area, including Mid Staffs. Published as a monograph in 2008, it appeared in a shorter form as a paper in the *BMJ* in 2009.³

Its conclusions were that HSMRs are unfit for comparing quality of care between hospitals because two of the measures used to make adjustments for different case mixes—the Charlson index, which adjusts for comorbidities, and the proportion of emergency admissions—actually increase the biases they are meant to reduce. “Claims that variations in HSMR from the Dr Foster Unit reflect differences in quality of care are less than credible,” the *BMJ* paper concluded. The full report said: “We find little evidence to support a systematic link between the Dr Foster SMR and quality of care or organisational failure,” a conclusion that certainly suited Mid Staffs, which was defending its clinical record in the face of HSMR figures consistently above the England average.

Jarman does not claim that the Birmingham researchers tailored their results to suit their paymasters, but he does suggest that they were asked to conduct the study only because of their long standing and well known scepticism about HSMRs. He argues that the conclusions they drew were wrong and that those errors could have been avoided if they had sought clarification from the Dr Foster Unit at Imperial College London, which he heads. Lilford, who did not give evidence to Francis, told the *BMJ* that he stands by the conclusions of the paper and has since elaborated on them in further work, including a recent paper in *BMJ Quality and Safety*.⁴

His view is that HSMRs are a useful test only if a substantial proportion of hospital deaths are preventable. “If less than 10%

Hospital standardised mortality ratios

The hospital standardised mortality ratio (HSMR) is the actual number of deaths at a hospital divided by the expected number of deaths over the same period, multiplied by 100. Each year the average HSMR score for all English hospitals is rebased at 100, so hospitals with a score of less than 100 have fewer deaths than expected and those with a score of more than 100 have more than expected. Scores are published annually in the *Hospital Guide* by Dr Foster Intelligence, and typically the range runs from about 75 to about 120.²

Actual deaths should be easy to measure. But calculating expected deaths requires some clever data management to account for factors that may differ from hospital to hospital, such as age, sex, diagnosis, ethnicity, admission source, type and length of stay, deprivation, and how ill patients are (comorbidity). The data comes from Hospital Episode Statistics supplied by hospitals as administrative data, based on codes that are a shorthand for each patient's diagnosis. The diagnoses used cover 80% of deaths in English hospitals. Comorbidities are measured by the Charlson index, which predicts the mortality for a range of conditions.

The method was developed by Brian Jarman and commercialised by Dr Foster, set up in the 1990s as a company specialising in health data analysis. Jarman continues to work at Imperial College in the Dr Foster Unit, which is part funded by the earnings of Dr Foster Intelligence. Other mortality indices exist, such as the risk adjusted mortality index used by CHKS (originally launched in 1989 as Caspe Healthcare Knowledge Systems, and now part of Capita).

of deaths are preventable, it's a very bad test," he said. "It's hard to find any signal in the noise. So unless the proportion of preventable deaths is very much higher than most people think is plausible, it's very unlikely that mortality ratios are going to be a good measure of the quality of care."

He is unhappy about Jarman's suggestion that the study was commissioned as a smokescreen. "I would be very concerned if I thought anybody was saying anything derogatory about Rashmi Shukla," he said. "She simply wanted to know what was going on here. So if anybody tried to impugn the SHA for commissioning the work or suggest it was done as a cover-up, I'd be mad as a snake. It's possible to be anti-HSMR and still do a perfectly valid analysis. If all Brian Jarman is saying is that he disagrees with me, I'm not the slightest bit annoyed. If he wants to criticise the work, criticise the work. I don't mind that."

Jarman told the inquiry that it was reasonable of the SHA to seek an opinion, but "whether you would take an opinion from a group that was known to be critical of the HSMR is another thing. I would personally, if I had been working for the SHA, have taken an independent opinion."

False reassurance

It is clear that the Birmingham work was used by managers at Mid-Staffs to reassure themselves. The earlier independent inquiry (also chaired by Francis) concluded: "The University of Birmingham reports, though probably well-intentioned, were distractions. They used the Mid Staffordshire issue as a context for discrediting the Dr Foster methodology." Roger Davidson, head of external affairs at the Healthcare Commission at the time, said in evidence to the public inquiry that the publication in the *BMJ*, on the day the commission planned to publish an investigation into Mid Staffs, "looked to me like a planned attempt at a spoiler."

Steve Allen, head of planning and information at the SHA, told the inquiry that he had been unable to make sense of claims from the commission that death rates at Mid Staffs were unduly high. He said: "I was also concerned that they appeared not to have given consideration to what I believed were compelling arguments presented by Dr Mohammed on behalf of Mid Staffordshire on why the apparently high death rate for emergency admissions was the result of a simple artefact—ie, that uniquely Mid Staffordshire did not submit 'short length-of-stay' emergency episodes to the national hospital episode statistics system which both Dr Foster and HCC [Healthcare Commission] used for their analysis."

In fact, Jarman said to the inquiry, including or excluding short stay patients makes a trivial difference to the HSMR, and the inclusion of comorbidities by using the Charlson index makes a change of less than 3%—in favour of Mid Staffs. So, in his

view, neither of the criticisms raised by Mohammed and Lilford is valid. Since it is now accepted that death rates at Mid Staffs were unduly high, it is clear in retrospect that clinicians and managers at the trust should have taken more notice of what the mortality figures were telling them.

Coding changes

Instead, they were persuaded by the Birmingham work that the problem was one of data quality, not patient care, and that the difference between Mid Staffs and other similar hospitals could be explained by differences in the way patients' conditions were recorded. This set the scene for the second and more serious charge levelled by Jarman at Mid Staffs and the SHA, that they manipulated patient coding to flatter their HSMRs.

There was certainly an abrupt change, principally in the use of the palliative care code Z51.5, which rose at Mid Staffs from below the England average in the final quarter of 2007 to well above it by the second quarter of 2008. The use of the code was increasing everywhere as a result of a change in the coding rules in March 2007, but Mid Staffs, Walsall, and George Eliot (together with Medway, a hospital outside the area) showed much larger increases than the average.

At Mid Staffs the proportion of deaths so coded rose from 0% (just one patient) in the fourth quarter of 2007 to 34% (72 patients) in the third quarter of 2008, while the trust's HSMR fell from 116 (above average) to 86 (below average) between the first and third quarters of 2008. So a dramatic rise in palliative care coding was followed three months later by an equally dramatic improvement in HSMRs. At the same time, a declining proportion of patients who died after fracturing a hip had this condition classified as their primary diagnosis, another change that would have the effect of flattering the trust's HSMR. Was it merely a coincidence that these changes in coding practice coincided with the launch of a Healthcare Commission investigation into Mid-Staffs in March 2008? Jarman told the inquiry: "I cannot comment on, because I do not know, the underlying reason for the change of palliative care coding and primary and secondary diagnoses at Mid Staffs (and also at George Eliot and Walsall)."

Sandra Haynes-Kirkbright, appointed data quality manager at Mid Staffs in July 2007 and responsible for overhauling the hospital's coding department, says that it was a coincidence that the changes coincided with the commission's investigation. "It is categorically not the case that palliative care is being used as a tool to disguise or hide deaths," she told the inquiry. Nor was the change in coding for fractured neck of femur "in any sense a fiddle," she insisted.

She explained that at Mid Staffs patients admitted with fractured neck of femur tend to be treated for the broken bone quickly, and rehabilitation begins. In some cases the patient fails to

mobilise sufficiently and subsequently contracts another condition, such as pneumonia. “According to the rules of coding I am obliged to code pneumonia as the first diagnosis,” she said. “The fact that fractured neck of femur changed from being one of the highest causes of death to one of the lesser causes merely reflects the fact that the coding was improved to meet national standards.”

Did Mid Staffs and the SHA know that these coding changes, whether introduced innocently or with deliberate intent, would make their mortality ratings look better? There is evidence they did but, like everything else in this tangled tale, it is contested. Dr Foster was not alone in using routine NHS administrative data to draw conclusions about clinical quality, and a rival company, CHKS, was also consulted by Mid Staffs.

Phillip Coates, a diabetes consultant at the trust, told the inquiry: “We had our data analysed by CHKS, which is another of the data manipulators in the field, who suggested to us that we did not have a mortality problem. And I think that gave us inappropriate and false reassurance. We knew that certain other trusts nearby used the CHKS system and they had similarly a less than satisfactory Dr Foster’s (sic) report, where the CHKS report has said exactly the opposite. . . . I think we were misled by the alternative analysis by CHKS.”

CHKS denies that it ever undertook such an analysis for Mid Staffs. It told the inquiry that it had undertaken a review of coding procedures in late 2006 because the trust feared it was not getting the full income to which it was entitled under payment by results. CHKS made a number of recommendations, including appointing a new head of coding (the post to which Haynes-Kirkbright was appointed) but did not, it says, undertake any mortality analysis.

Plainly, however, this is not the view taken at the time, or retrospectively, by the trust and its managers. In a report in April 2008 by the trust’s chief executive, Martin Yeates, to local commissioning bodies, he wrote: “An independent review of our data capture and coding was undertaken by the organisation CHKS who confirmed that our level of data capture, particularly with regard to comorbidity of patients and the coding of this activity were well below average. Most importantly, they also confirmed that our overall mortality rates were in line with national averages as would be expected at this type of organisation providing our type of care.”

And in a response to CHKS’s denials, Coates said that three of the trust’s senior team—himself, Yeates, and Valerie Suarez, medical director of the trust—had made the same comments about being reassured by CHKS. “As there appears to be no documentation from CHKS that such a calculation was ever made, I can only suggest that we (Martin Yeates, Dr Suarez,

and myself) were either given this information informally during a conversation with representatives of CHKS or we were all mistaken in our understanding,” he told the inquiry.

CHKS has certainly commented on mortality rates at other hospitals, including Medway, where its advice on use of the palliative care code resulted in an increase from 8% to 37% in the proportion of deaths so coded, and a corresponding fall in the hospital’s HSMR.⁵ In evidence to the inquiry, however, Paul Robinson of CHKS said the company had not advised Mid Staffs about mortality since the 1990s.

There is no evidence to support a claim that the change of coding practice at Mid Staffs, Walsall, and George Eliot was a concerted effort orchestrated by the SHA. In an exchange of emails with Jarman shown as evidence to the inquiry, Mike Brown, the medical director of Walsall, acknowledges that his hospital “got it wrong” but says that the problems were “entirely internal, no shared work with any other agency or organisation that I am aware of.”

In the broader context of failings at Mid Staffs, the row over HSMRs is fascinating but something of a sideshow. It illustrates both the strength of mortality measures, in that they consistently provided a signal that something was seriously wrong, and their weakness to changes in coding practice, however that comes about.

The new NHS now has its own version of HSMRs, called summary hospital level mortality indicators (SHMIs). The indicator has incorporated some changes, such as including all deaths occurring within 30 days of hospital treatment instead of only in-hospital deaths, but is essentially the same measure. Will it prove impervious to gaming? It would be a brave man who would bet on it.

Competing interests: I have read and understood the BMJ Group policy on declaration of interests and have no relevant interests to declare.

Provenance and peer review: Commissioned; not externally peer reviewed.

- 1 Kark T. *Closing submissions of counsel to the inquiry*, 9 December 2011, www.midstaffspublicinquiry.com/node/505.
- 2 Dr Foster Intelligence. *The hospital guide 2012*. <http://drfosterintelligence.co.uk/thought-leadership/hospital-guide/>.
- 3 Mohammed MA, Deeks JJ, Girling A, Rudge G, Carmalt M, Stevens AJ, et al. Evidence of methodological bias in hospital standardised mortality ratios. *BMJ* 2009;338:b780.
- 4 Girling AJ, Hofer TP, Wu J, Chilton PJ, Nicholl JP, Mohammed MA, et al. Case mix adjusted hospital mortality is poor proxy for preventable mortality: a modelling study. *BMJ Qual Saf* 2012;21:1052-6.
- 5 Robinson P. Why end of life care should be recorded. *Health Serv J* 2008 Dec11. www.hsj.co.uk/resource-centre/best-practice/why-end-of-life-care-should-be-recorded/1936345.article.

Cite this as: *BMJ* 2013;346:f562

© BMJ Publishing Group Ltd 2013