Clinical governance: expect not the best but the ideal

I have, as I've said before, led a busy and satisfying life in medicine, but as I near the end of my consultant career, I have begun to wonder whether medical politicians, even practitioners, could learn from other walks of life. This became particularly pertinent when I was asked to advise the Health Minister on how audit data should be used in clinical governance. Having recently renewed contact with Charles, one of my contemporaries at university and, as I discovered at dinner a few weeks' earlier, an ostensibly conservative figure but a radical thinker, I did not hesitate to call him and tell him my fears.

---

'Charles, I'm worried about how audit data might dictate our clinical practice. How do you allow for differences between patients? If I maintain the outliers patients standards, how can I be sure that my consultations are not compromised by hurrying? If you were advising the Health Minister, what would you say?'

'Well, Coe, I'm not too concerned about your first point. Differences may be exaggerated and it should be possible statistically to discount them. The implications of your second worry do, however, concern me – and don't forget the importance of chance in explaining differences. I suggest you begin by giving the Minister a good illustration of the fact that 100% (or 0%) is unlikely to be an ideal target. For example, 0% operative mortality is more likely to indicate a timid approach than a brilliant anaesthetist, while 100% participation in clinical trials almost certainly indicates that patients have been pressed to participate. I would tell the Minister that he must accept that everyone is trying to achieve the ideal. Therefore, until proved otherwise, the average performance must be regarded as the ideal. There should be no prejudice in considering deviation to either side as "good" or "bad".'

'But even I can see that he would accuse you of complacency', I said.

'He probably would, but let the argument develop. First plot the data, then look at the distribution. A normal (symmetrical) distribution that is narrow, suggests all units are maintaining an easily achieved consensus. A broad distribution would suggest disagreement, lack of enthusiasm or poor resources for some. You have to pay particular attention to the tails of asymmetrical distributions. A tail one way might suggest that a few are making worthwhile advances, that there has been over-allocation of resources, or that the majority are underperforming. The other direction might suggest particular difficulties, underprovision of resources or poor performance by a minority.'

'But that still means accepting the status quo for the majority.'

Charles continued: 'Yes initially, but successive audits will give you a picture of the movement of the median. This will provide you with a slow but effective method of defining demand for clinical services and levels of response to it; this will become increasingly informative as the implications of the asymmetrical tails become clearer. As external standards (preferably evidence-based) are developed, they should be entered onto the plot, providing a specific point at which to aim. If you then compare the median and standard, you will be able to set targets in the light of the separation between them and the shape of the distribution curve. These targets should be absolute, eg within 5% of the agreed standard (the ideal), rather than relative, eg the top quartile (the best 25%). By studying changes in the relationship between standard and median, and of the shape of the curve, successive audits will come to reflect the practicability, or perceived desirability, of the change. Rapid moves to narrow distribution about the standard would suggest desirability and practicability, whilst persistent broad distribution might imply disagreement, and the development of a tail, particular difficulties for some.'

'What about bad performers?' I asked, anxiously.

'Well, Coe, I would tell the Minister that these audits are bad at detecting poor individual performance because of the effects of chance. If 100 equally competent surgeons performed 100 operations with expected net mortality of 5%, what would you expect the range to be?'

'About two or three to eight percent,' I guessed.

'No, more like nought to fourteen percent,' he replied, 'Most familiar statistics assume an actual or putative sample. For example, power calculation involves estimating the size of the required sample. Here the whole population is necessarily observed, and confidence intervals must encompass everyone. So, 0.1% intervals must be applied when there are 250 trusts. This doesn't mean that less extreme outliers should not examine their own performance or have to examine it to the purchasers or the NHS Executive; it does mean, however, that where league tables are published, correct confidence intervals should be given and their implications fully explained. Poor performers are easier to identify when there is a clearly defined objective standard with a long tail in the perceived bad direction. Nevertheless, I would warn the Minister that he publishes at peril lest apparently poor performance, determined by chance, should become self-perpetuating, destroying unlucky but otherwise adequate units by depriving them of custom, and so on. The practicalities demand self collection of data, whilst the fear of unfair disclosure could increase the risk of fraud...'.

'Yes, I've already had wind of that.'

Charles paused before continuing: 'I think I would end, "Minister, forget about best and worst and let audit do what it should do: define the ideal and encourage its universal achievement").

I only wish I had the confidence that something like this will happen. One can only hope.'