Gwyn Bevan November 26th, 2024

Hospital league tables – an idea whose time has passed?

The Labour Government has announced the reintroduction of league tables as a way of measuring the performance of NHS hospitals. ***Gwyn Bevan*** *warns that the last time such league tables were introduced under the Blair Government lead to very mixed results, and points out the weaknesses of such an approach to incentivising healthcare providers.*

Enjoying this post? Then sign up to our newsletter and receive a weekly roundup of all our articles.

On November 13, 2024 the Department of Health and Social Care announced a package of tough NHS reforms with zero tolerance for failure. This regime will use a league table of performance of NHS providers, to sack persistently failing managers, send turnaround teams into struggling hospitals, and reward the best performers with earned autonomy. Its overriding objective is to "cut waiting times from 18 months to 18 weeks". This regime combines "targets and terror" and "naming and shaming". It looks like the annual star ratings, which was implemented in the NHS in England from 2000 to 2005, and enabled the reduction in hospital waiting times from 18 months in 2000, to 18 weeks by 2008. Alan Milburn, who was appointed on November 9 as "the Former Health Secretary to help government fix health and care", introduced the star rating regime into the NHS in 2000. So, what is the evidence that star rating regime was effective? What are the weaknesses of the new regime of zero tolerance for failure? And what changes are needed for the NHS to deliver high performance?

The vision and context of Blair's NHS rating system

Implementation of the star rating regime followed the "most expensive breakfast in British history". Tony Blair's interview on Breakfast with Frost, on 16 January 2000, was in the midst of winter NHS crisis of 1999/2000. That day he committed his government to increase real spending on the NHS by 5 per cent a year for five years with the objective of raising the percentage of GDP that the UK

spent on health care to the European average. Figure 1 shows the consequences. Figure 1 also shows why the NHS has been struggling after a decade of austerity. (The blip in 2020 was driven by the 10 per cent fall in GDP during the Johnson government's response to Covid – not increases in real spending on the NHS.)

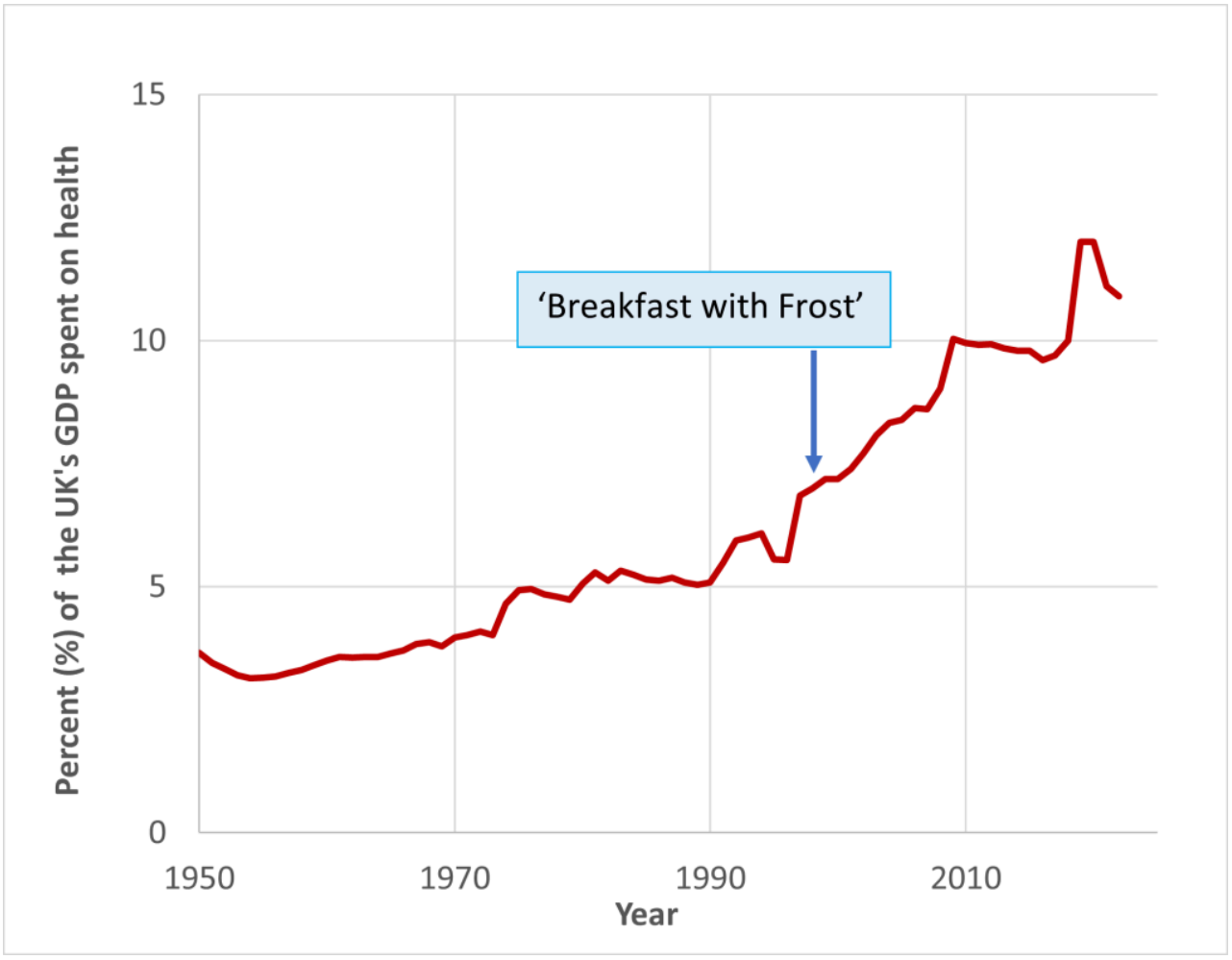

Figure 1: NHS spend as a per cent of the UK's GDP (from 1950 to 2022). Source: How Did Britain Come to This?

Alan Milburn won support from representative organisations of frontline NHS staff and patients for the principles of the NHS Plan of July 2000. That set out the rationale and policy of what became the star rating regime (called "traffic lights"). Its themes were that the failure of the NHS to modernise had been tolerated because of consistent underfunding; and the new funding settlement changed that and required ambitious targets for short waiting times. The devolved governments of Wales and Scotland also increased substantially NHS funding and aimed to reduce waiting times.

When NHS trusts failed to deliver the targeted reductions in waiting times, only the government in England changed from rewarding failure to imposing sanctions.

The evidence for the effectiveness of the star rating regime comes from the natural experiment following devolution in the UK from 1999. When NHS trusts failed to deliver the targeted reductions in waiting times, only the government in England changed from rewarding failure to imposing sanctions. The governments of Scotland, Wales and Northern Ireland continued to reward failure. A Nuffield Study comparing the four nations showed England's superior performance in reducing NHS waiting times. That finding for elective admissions was confirmed by rigorous econometric analyses comparing England with Wales and Scotland. A damning report by the Auditor General for Wales pointed out that, in 2005, the sum of the two waiting time targets to be referred to a specialist *and* admitted for an operation was nine months in England, and three years in Wales!

Three weaknesses of the regime of zero tolerance for failure

Finance, waiting times and scandals

The first weakness is the lack of resilience in England's systems of governance to prevent hospital scandals when the overriding priorities of the Department of Health are finance and waiting times. That was the common problem identified by the Kennedy Report and the Francis Report of the Public Inquiries into NHS scandals that developed under the different systems of governance under Margaret Thatcher, in the 1980s and 1990s, and Tony Blair in the 2000s. Parents of 160 brain-damaged children who had undergone paediatric cardiac surgery at Bristol called for a public inquiry; and the Kennedy Report estimated there had been 30 excess deaths there. The Francis Report described how "between 2005 and 2008 conditions of appalling care were able to flourish in the main hospital serving the people of Stafford and its surrounding area". The reason why was finance was a cause of the scandal at Mid-Staffordshire, when real spending on the NHS was increasing at 5 per cent each year, was that the Board's bid for earned autonomy (as an NHS Foundation Trust) depended on eliminating its financial deficit. So, the Board made severe cuts to its nursing staff resulting in appalling care for years, which the Francis Report found that was hidden in plain sight from "a plethora of agencies, scrutiny groups, commissioners, regulators and professional bodies".

The second weakness of the proposed league table is its dependence on measuring the performance of each acute hospital as a single score

The new regime of zero tolerance is being implemented without the largesse of the 2000s: The Health Foundation estimates that the Autumn Budget will increase NHS expenditure in real terms by 3 per cent between 2023/24 and 2025/26. In the 2000s, the UK enjoyed sustained economic growth, which is now uncertain. NHS England reported that, in July 2024, 31 out of 42 of England's integrated care systems were in deficit (total £2.2bn). And that it was imperative that all achieved their plans to eliminate their deficits by March 2025. Last month, the Dash Report on the Review of the Care Quality Commission found "significant failings" in the internal workings of CQC, which have led to a substantial loss of credibility within the health and social care sectors, a deterioration in the ability of CQC to identify poor performance and support a drive to improve quality. This combination has all the makings of another NHS scandal.

The reductionism of a single score system

The second weakness of the proposed league table is its dependence on measuring the performance of each acute hospital as a single score (as did the star ratings). Assessments by CQC show that performance varies substantially within each hospital. So, crude, reductionist rankings cast an unjustified shadow on those delivering a high quality service in a hospital with a poor ranking in aggregate, and vice-versa.

A wonderful example of an effective learning organisation is from how all the districts in the region of Tuscany learnt from the best practice of diabetic care in Arezzo.

By contrast, Italy's Inter Regional Performance Evaluation System (IRPES) uses the famous PISA dartboard to display performance across multiple dimensions (as shown in Figure 2 for the regions of Marche and Tuscany). Indicators with excellent performance are in the green zones near the centre of the dartboard; those with poor performance are in the red zone on the outer circle. Hence it is clear that performance of Tuscany was superior to that of Marche. Performance on the ultimate goal of the health of the local population is reported above the dartboard.

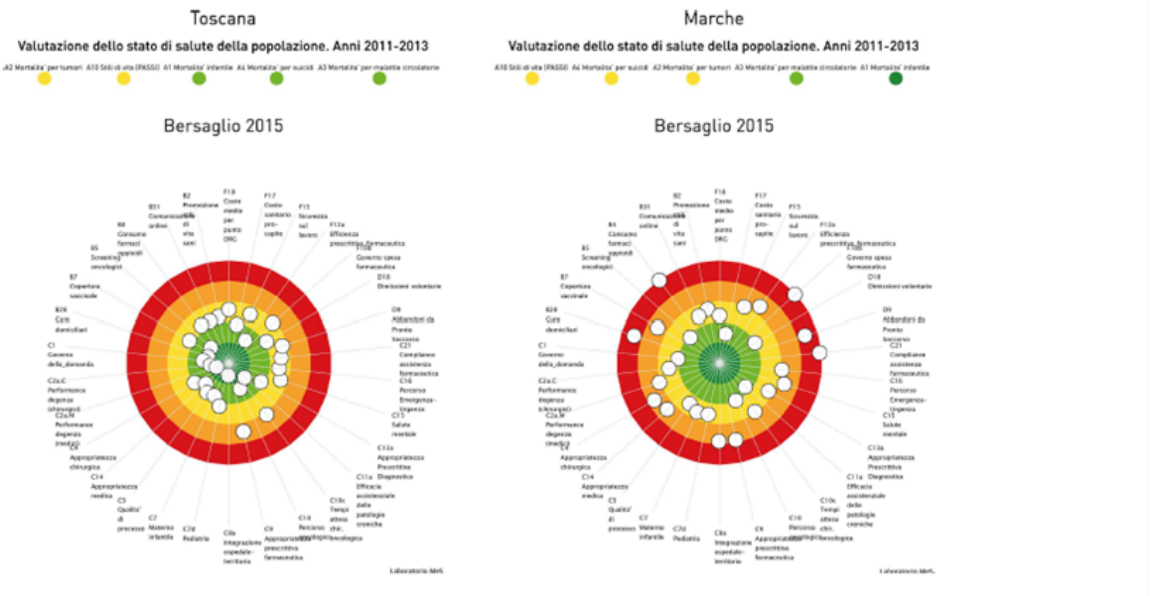

Figure 2: The Tuscan dartboards displaying Performance for Regions of Tuscany and Marche . Source: Reputations count: why benchmarking performance is improving health care across the world

Undermining the capability of the NHS to learn

The third weakness is using the proposed league table for both trusts deemed to be "failing" *and "*high performing" (as in the regime of star ratings). A system that rewards those deemed to be high performing with "earned autonomy" is antithetical to developing a learning organisation within the NHS. As Adam Oliver has argued, managing performance of public services where markets fail requires governance with two different systems: of sanctions limited to unacceptably poor performance organised at the national level, and a collegial system of learning from best practice informed by benchmarking in regional networks, as in Italy's IRPES. A wonderful example of an effective learning organisation is from how all the districts in the region of Tuscany learnt from the best practice of diabetic care in Arezzo.

What changes are needed for the NHS to deliver high performance?

Three radical changes are required if the NHS is to develop the capability to deliver high performance:

- Return to the quaint tradition of ministerial accountability, which Aneurin Bevan saw as fundamental to the governance of the NHS so “echoes would reverberate throughout Whitehall every time a maid kicked over a bucket in a hospital ward”.
- Stop requiring CQC to do mission impossible by assessing performance within each NHS acute hospital, and create a NHS regulator with the limited remit of inspecting whether each system of governance is designed to prevent unacceptably poor performance.
- Create networks of regional governance using benchmarking to learn how to improve performance.

Image credit: chrisdorney Shutterstock

About the author

Gwyn Bevan

Gwyn Bevan is Emeritus Professor of Policy Analysis in the Department of Management at LSE and author of How did Britain come to this? A century of systemic failures of governance. He was Director of the Office for Health Care Performance at the Commission for Health Improvement, which regulated quality of care in the NHS in England and Wales from 2000 to 2004.

Posted In: Government | LSE Comment