The Mid Staffordshire dilemma: Not a matter of choice

> *The backdrop facing all healthcare systems is that the costs of providing treatments are rising faster than our ability to pay for them."*

The Francis Report into failings at Mid Staffordshire Hospital Trust sheds light on a widespread and potentially catastrophic dilemma facing leaders in the NHS: the conflict between achieving financial sustainability on the one hand and delivering high quality and timely care on the other.

This article explores how hospital systems can spiral into chaos when dangerous choices to save costs lead to a loss of control of the quality and timeliness of patient care and why management and clinical attention can be diverted from these over-riding objectives. It also explores how patient-centric systems to increase patient flow by reducing delay can not only rapidly improve patient outcomes but also reduce waste and transform financial performance using principles derived from the Theory of Constraints (TOC).

The backdrop facing all healthcare systems is that the costs of providing treatments are rising faster than our ability to pay for them. Faced with increasingly onerous mandates to control costs, financial performance is easy to see and measure and receives more attention. Patient safety is much harder to assess until a catastrophe happens.

In Mid Staffordshire's bid to achieve foundation status, a cornerstone of the country's current health strategy, it chose to reorganise and reduce frontline staff. This decision set the trust on a turbulent path that would eventually make the system unmanageable and lead to unnecessary patient deaths.

Seemingly complex systems

Systems and processes involved in delivering healthcare seem complex and involve the work of many people distributed in time and space. Much of this complexity appears dynamic and difficult to see, with actions and decisions made in one part of the system creating consequences elsewhere and at a later time, seemingly unconnected. This makes it difficult to focus attention on those areas that will make a real difference.

Hospitals have a finite number of managers and clinicians and they can address the organisation and patients needs for a limited amount of hours in the day. As a result the managerial and clinical capacity available for giving attention to addressing the organisations needs is finite. The Mid Staffs example is just one of many examples where there are plenty of things that require attending to. Like many organisations it is evident that management and clinical attention is itself a bottleneck whose demand exceeds the capacity available.

To understand the significance of this in the Mid Staffs example it is important for us to understand how this limited capacity is used to improve the performance of the hospital. There is always a risk that much of this capacity is often wasted and sometimes not just wasted but diverted and used in a way that can actually cause a set back. A deeper examination of the broader health sector shows how we have created elaborate and complicated systems and measures that do just that. They waste management and clinical attention and in many cases do not just waste it but divert attention towards activities that do damage rather than improve the system. Dr Goldratt, the inventor of the Theory of Constraints highlights that this seemingly grotesque situation must stem from a more fundamental human behaviour namely:

"The managerial and clinical capacity available for giving attention to addressing the organisations needs is finite. It is evident that management and clinical attention is itself a bottleneck whose demand exceeds the capacity available."

1. Our fear of complex systems that drives us to dissect these seemingly complex systems into sub systems leading to divert management and clinical attention to chose local optima which are not in line with the global objectives.

2. Our fear of the unknown that drives us to finer and finer resolution going into more and more detail that averts management and clinical attention to optimise within the noise.

3. Our fear that conflict will lead to a tug of war both between and within groups of managers and clinicians that divert attention to constantly struggle with unacceptable compromises.

Threats to performance

In our experience of working in acute NHS hospitals, the biggest threat to performance is the amount of time doctors, nurses and managers spend dealing with disruption and delay rather than providing high quality and timely care. For patients, this means much of their journey is spent waiting for treatment rather than receiving it. When each individual delay is multiplied by

the thousands of patients admitted each year, the enormous scale of waste in productive capacity and operating expense felt by the entire healthcare community, including acute, community, rehabilitation and mental health care, comes into stark focus.

This disruption and delay is evidenced through queues that come and go throughout the system – the phenomenon of the wandering bottleneck – and is most often due to a lack of prioritisation or synchronisation in the delivery of treatment. The outcome is that many patients will be rehabilitated more slowly than necessary, often remaining in expensive acute settings longer than they need to, increasing the cost of their treatment and the chance of them succumbing to secondary health complications, particularly in the case of elderly patients. Reducing capacity through cuts to staff and resources will quickly create queues and intensify delays, forcing patients to wait longer than they should for treatment.

"For patients much of their journey is spent waiting for treatment rather than receiving it. Reducing capacity through cuts to staff and resources will quickly create queues and intensify delays, forcing patients to wait longer than they should for treatment."

Extensive research and common anecdotal evidence from health professionals indicates that the quality and timeliness of care rapidly deteriorate when staff are overstretched or have too many simultaneous demands on their time. Indeed, catastrophic failures most often occur during extended periods of unreasonable staff pressure. At these times, systems can shift from being predictable and stable, sustained by staff working harder to mask issues, to suddenly becoming unmanageable and chaotic. This transition is non linear and all too often the hospital will rapidly slip from a coping strategy into one of catastrophic failure.

Patient-centric care

Theory of Constraints (TOC) principles assert that the more complex a goal oriented system appears, the fewer constraints there are impacting on its performance. However, the repercussions of these constraints will be far-reaching and impede the performance of every part of the system. To identify these few underlying constraints hospitals need to identify which resource or task resource combination causes the most disruption and delay to most patients' journeys. This is where improvements have the most positive system-wide impact on patient flow, they result in the quickest and most dramatic improvements in performance and avoid wasting efforts on improving non-constraining resources. This has a profound impact on the underlying dilemma as it gives us, for the first time, the opportunity to improve the system at a faster rate than rising medical costs.

To achieve such improvements requires clinicians to evaluate each patient's planned discharge date based on realistic estimates of the time it will take them to be fit for discharge. A plan of key clinical and administrative tasks to be performed within this time is also agreed. Within 24 hours of admission, each patient is given a project plan and planned date of discharge. These are recorded in an active database that all clinicians, managers and referrers have access to. This level of transparency helps staff to identify potential problems before they happen and address them so that they don't delay a patient's planned discharge.

As well as making considerable improvements to the day-to-day delivery of care and patient outcomes and experiences, each hospital has also improved the quality and timeliness of care to increase the rate at which patients flow through their hospitals, all within budgets and without increasing capacity. Staff confidence, productivity and job satisfaction have also been shown to improve as clinical and support staff are given realistic and achievable priorities, and a strong decision-making culture is now palpable on the wards.

"We need to focus our attention on exploring and evaluating new ways that shed new light in the surprising, realistic potential for improving the overall performance of our systems for delivering healthcare."

Threats, mandates and targets

The demands for more stringent policing and punishments for hospitals and their leaders who do not meet minimum standards suggested as resolutions to the current crisis facing the NHS will not work in organisations like Mid Staffordshire Trust. Shouting louder and making ever-increasing demands usually sends people into a spin and makes it even less likely that they will ever recognise the true reason for the system failures.

There is a considerable risk that many of the recommendations made so far to prevent this crisis from happening again are being misinterpreted. Creating lists of people who need to be fired in the wake of the Mid Staffordshire case, and legislating to make managers personally and criminally responsible for tragedies ensuing from inefficient practice will only deter more talented and visionary professionals from entering and leading our healthcare system out of its difficulties. Instead we need to focus our attention on exploring and evaluating new ways that shed new light in the surprising, realistic potential for improving the overall performance of our systems for delivering healthcare.

Mike Williams is a former NHS CEO and now senior research fellow at the University of Exeter Business School and Non-Executive Director of Plymouth Hospitals NHS Trust and Alex Knight is founding partner of QFI Consulting, which is a leading specialist in Theory of Constraints.