

Managing capacity and demand across the patient journey

Steven Alder, Kate Silvester and Paul Walley

ABSTRACT – Bed availability remains the main operational focus for managers and clinicians on a day-to-day basis within the NHS. There is now published research that establishes a lack of bed stock is too simplistic an explanation of the situation. Other reasons for bed shortage include the daily and weekly lack of synchronisation of admissions and discharges, the large variation in bed occupancy over time, the downtime during weekends and holiday periods, wasted time during inpatient stays and the variation in patient length of stay. So far most of what little work has been done has focused on the front end of the process, to ‘buffer’ unplanned admissions through the use of short-stay facilities, such as medical assessment units, as a short-term solution. This paper reviews the evidence for the hypothesis that bed availability problems can be solved by actions other than the addition of more beds to the system.

KEY WORDS: accident and emergency flow, admissions and discharges, bed management, synchronisation, systems planning, variation

Introduction

For two decades at least there has been concern that the NHS has been reducing acute bed capacity at too fast a rate to cope with demand for healthcare. The NHS Plan set out a significant expansion of the bed capacity of the system as a whole, although not all of this expansion was within the NHS and relies on external providers for some capacity.¹ The extent to which this plan has been implemented to date is debatable.² Earlier work has explored some of the capacity and demand variation issues across the NHS, suggesting that more capacity is not necessarily the answer to waits and delays for care.³ In particular, lack of bed availability is, from our perspective, more a symptom of significant planning defects and system design issues.

The purpose of this paper is to explore some of the conflicting views concerning the need for more acute bed capacity and to establish evidence to support a different approach to managing a healthcare system. In this paper a ‘systems’ perspective is taken and the impact of variation on the patient journey through acute care and onto intermediate or community-based care is considered. Work is underway to establish methodologies for seeking solutions to these problems.

Steven Alder, consultant neurologist, Department of Neurology, Plymouth Hospitals Trust; **Kate Silvester**, Inquiry into Flow, Cost and Quality Programme lead, The Health Foundation c/o South Warwickshire General Hospitals NHS Trust, Warwick Hospital; **Paul Walley**, associate professor, Warwick Business School, Warwick University

Bed management as a system issue

Bed availability is clearly a key issue in all acute hospitals. The four-hour government target for the accident and emergency (A&E) journey time put special focus on emergency medical admissions. It is now clear that bed occupancy levels influence the ability of A&E departments and assessment units to send patients to specialist wards in a timely manner.⁴

Bed management is the typical countermeasure to the almost daily ‘bed crisis’.^{5,6} Black criticises bed management in a typical hospital, especially concerning how information is managed: ‘...hospitals typically operate like a hotel where the only way to find a free room is to knock on every door.’⁷ However, some believe that bed shortages are self-inflicted and compounded by the known, predictable system dynamics that create additional problems for bed availability.^{6,8}

The daily bed cycle

In the majority of NHS hospitals, the peak in arrivals (especially for elective patients) occurs during the morning, but the peak in discharges occurs in the late afternoon. One of the key roles for short-stay or assessment units is that of an operational ‘buffer’, where patients are kept during daytime hours, while the system waits for the later afternoon and early evening discharges to occur. Simple calculations can be made to demonstrate that typically 10% of bed capacity is tied up in managing this problem.

Weekly cyclical patterns

There is evidence for a weekly variation in admissions and discharges. First, the peak in arrivals of emergency admissions occurs on a Monday – precisely the day of the week when most elective admissions are planned. Second, discharge patterns are not synchronised with demand and do not usually occur seven days a week in the medical specialties. A telling symptom of this problem is that patient length of stay (LoS) varies depending on which day of the week they are admitted. It has been estimated that large output increases could be achieved from existing bed stock if admissions and, particularly, the discharges were managed seven days a week.⁹

Annual cyclical patterns

Healthcare demand is surprisingly non-seasonal when compared with demand in other parts of the service economy. It is too simplistic to suggest that ‘winter bed pressures’ are purely caused by increases in demand at this time of year. There are other complicating factors, such as the impact of lack of decision

making and discharge capacity due to holiday leave. The annual cycle of capacity availability, where activity falls at key times of year (Christmas, the week of the February school half-term, August), contributes to high bed occupancy.

The flaw in long-term planning

A key planning issue for many years has been the failure to coordinate the impact of random variation of demand (both patient numbers and LoS) and availability of processing capacity with the staff on bed requirements. If bed requirements are based on average demand and average length of stay then there will not be sufficient beds. As soon as there is a random variation in demand and a random variation in staff capacity, bed shortages will occur on a frequent basis due to 'Erlang' queues.¹⁰ A number of studies have shown the impact of variation on the need for less than 100% occupancy of beds.^{11,12} Bagust *et al*, modelled a 200-bed emergency system and demonstrated that it could not safely operate at above 85% average occupancy and simultaneously maintain reasonable bed availability during peaks in demand.¹³ However, it is important that the findings in the paper are not misinterpreted. The assumptions made in this study mean that any generalisation has to be done with care. Key issues are:

- 1 The study is of an emergency system. In an elective-only system, where short waiting times for patients can be tolerated and arrivals can be scheduled, the occupancy of the system in theory could be higher.
- 2 The scale of the healthcare system influences achievable occupancy. As the bed pool is increased, the impact of demand variation becomes a less significant factor. Large systems might achieve 90% occupancy, whereas small units, eg an intensive care unit, in theory cannot expect to achieve anywhere near 85% to maintain availability.
- 3 The variation of LoS influences the model. Where LoS can be controlled, eg by reducing the variation of staff availability, utilisation can increase.
- 4 The occupancy results from the model are an expected average. In the real situation, clinicians or managers must not set 85% as some arbitrary target. The random variation over time in such a system means that the occupancy naturally varies between (say) 70% and 100%. More generally, there is the issue of whether bed state is self-regulating to some extent, as discharge effort, admission thresholds and elective admissions may vary naturally as a consequence of bed state.
- 5 Variation in the system is amplified along the care pathway due to known system effects.¹⁴ Improvement of flow through the system as a whole requires either increased relative capacity at the downstream end of the system (discharge or community care) to increase the probability of a discharge occurring or to reduce the variation at every step in the pathway.

It is the control of these sources of variation that form the main focus of suggestions for solutions to the problem of bed shortage.

Can we create smoother demand?

It is ironic that all relevant studies show that emergency admissions experience less significant variation than elective patient admissions to hospital.¹⁴ Clinical schedules distort the natural pattern of demand. Two aspects of this variation in clinical capacity need to be assessed. First, the daily fluctuations in capacity must be adjusted over a seven-day week where possible. In particular, the system for discharging patients at the weekends should be the same as on weekdays. Second, there is a tendency to batch patients by specific treatment. This increases the natural mix variation of what services are required on a daily basis and distorts demand for services downstream such as, radiology, pharmacy etc. As this impact is cumulative, there are severe consequences for inpatient delays (waiting for test results etc) that affect LoS.

Can we match capacity to demand through more flexible working?

One of the characteristics of healthcare is the relative lack of flexibility in work scheduling. Clinical schedules do not match demand. Perhaps of even greater concern is the lack of certainty of availability of staff, through random failure (staff absenteeism, multi-tasking etc). These faults prevent the system from being able to respond to short-term random variations in demand or match rarer seasonal peaks. This mismatch leads to backlogs of work that are 'stored' in beds.

Can we better manage patient length of stay?

There is some evidence that the UK hospitalises patients to a greater extent than comparable 'efficient' systems in other countries and that longer LoS are as significant as lower admission thresholds.^{15,16} Clark and Rosen usefully debate whether or not a shorter LoS is beneficial for patients concluding that some can benefit from shorter LoS.¹⁷ There is already some evidence to show that shorter stays can have beneficial effects on patient outcomes if the systems are improved in a sustainable manner.¹⁸

From the frequency distribution of LoS, it is recognised that patients cannot be considered as a single homogeneous group. The vast majority of emergency patients require a very short LoS (the quick patients), some patients have severe pathology that requires specialist support and time to heal, and finally there is a small but significant group that will never recover their original function and require extensive rehabilitation and support to get them home. Some of these will require long-term care or end-of-life care.

Using their average LoS as a measure of effective and efficient flow through the system as a whole is unhelpful. This encourages arbitrary bed reductions to force average LoS reductions that are not productive¹⁹ and care must be taken to ensure LoS patterns are sustainable.

The frequency distribution of the LoS and clinical experience suggest that these patients follow very different processes

with different dynamics. There are common and more specific reasons for delays in each group:

- Ward rounds are relatively infrequent in many trusts and this artificially adds days onto LoS, especially for patients who require a short stay. The problem is at its worst where patients do not recover sufficiently for discharge on a Friday and usually remain until Monday due to lack of weekend cover. Similar problems occur before discharge, when one visit to each patient per day by a decision maker is the norm. This causes simple activities, such as requesting, interpreting and responding to a simple test, to take several days when the total time needed is rarely more than one hour.
- In-hospital access to support services, such as physiotherapy, suffers from very similar variation issues and there is poor coordination of these services. Access issues add considerably to LoS.
- Junior staff perform much of the key work, creating information handover issues and delays to key decisions.
- For patients with complex discharge needs, availability of suitable non-acute support in the community will depend, in part, on the 'batch production' system in social and primary care services. There is also some evidence that structural changes in previous decades, to reduce community bed capacity in England, have had a significant impact on the ability to discharge patients into suitable long-term care.^{20,21}

Conclusions

The ability of hospital staff to get a patient to the right bed at the right time is dependent on bed occupancy. However, delays in the NHS are partly a self-inflicted problem due to a fundamental flaw in the NHS planning process whereby the impact of demand and capacity variation is not understood or managed. It is impossible to run a safe and efficient system in which there is unnecessary variability in the demand and capacity. In situations where resources are managed using average requests for beds and average LOS, delays to care will happen. A 'systems' approach is therefore recommended to understand that different patients require different processes due to their different dynamics. The whole health and social care system should plan capacity to meet the very predictable and relatively constant demand at every step in these processes, thus avoiding the need to use beds as a storage system for patients caught up in the lack of system flow.

References

- 1 Department of Health. *Shaping the future: long term planning for hospitals and related services. Consultation document on the findings of the National Bed Inquiry*. London: DH, 2000.

- 2 Macfarlane A, Godden S, Pollock AM. Are we on track – can we monitor bed targets in the NHS plan for England? *J Public Health* 2005;27:263–9.
- 3 Silvester K, Lendon R, Bevan H, Steyn R, Walley P. Reducing waiting times in the NHS: is lack of capacity the problem? *Clinician Manag* 2004;12:105–11.
- 4 Cooke MW, Wilson S, Hallsall J, Roalfe A. Total time in English accident and emergency departments is related to bed occupancy. *Emerg Med J* 2004;21:575–6.
- 5 Boaden R, Proudlove N, Wilson M. An exploratory study of bed management. *J Manag Med* 1999;13:234–50.
- 6 Proudlove NC, Gordon K, Boaden R. Can good bed management solve the overcrowding problem in accident and emergency departments? *Emerg Med J* 2003;19:49–55.
- 7 Black S. More and better management is the key to fixing the NHS. *BMJ* 2006;333:358
- 8 Fusco D, Saitto C, Arca M *et al*. Cyclic fluctuations in hospital bed occupancy in Roma (Italy): supply or demand driven? *Health Serv Manag Res* 2003;16:268–75.
- 9 Black S, Proudlove N. Hospital bed utilisation in the NHS and Kaiser Permanente, bed management in the NHS can be improved easily. *BMJ* 2004;328:582–3.
- 10 Noon CE, Hankins CT, Cote MJ. Understanding the impact of variation in the delivery of healthcare services. *J Health Manag* 2003;48:81–8.
- 11 Gallivan S, Utley M, Treasure T, Valencia O. Booked inpatient admissions and hospital capacity: mathematical modelling study. *BMJ* 2002;280–2.
- 12 De Bruin AM, van Rossum AC, Visser MC, Koole GM. Modeling the emergency cardiac in-patient flow. An application of queuing theory. *Health Manag Sci* 2007;10:125–37.
- 13 Bagust A, Place M, Posnett JW. Dynamics of bed use in accommodating emergency admissions: stochastic simulation model. *BMJ* 1999;319:155–8.
- 14 Walley P, Silvester K, Steyn R. Managing variation in demand: lessons from the UK National Health Service. *J Health Manag* 2006;51:309–20.
- 15 Ham C, York N, Sutcliffe S, Shaw R. Hospital bed utilisation in the NHS, Kaiser Permanente, and the US Medicare Programme: analysis of routine data. *BMJ* 2003;327:1257.
- 16 Jarman B, Aylin P, Bottle A. Discharge destination and length of stay: differences between US and English hospitals for people aged 65 and over. *BMJ* 2004;328:605.
- 17 Clarke A, Rosen R. Length of stay. How short should hospital care be? *Eu J Public Health* 2001;11:166–70.
- 18 Gilligan S, Walters M. Quality improvements in hospital flow may lead to a reduction in mortality. *Clin Govern Int J* 2008;13:26–34.
- 19 Rae B, Busby W, Millard PH. Fast-tracking acute hospital care – from bed crisis to bed crisis. *Aust Health Rev* 2007;31:50–62.
- 20 Millard PH. NHS needs plan for all acute, rehabilitative, and long-stay care. *BMJ* 2001;322:735.
- 21 McCoy D, Godden S, Pollock AM, Bianchessi C. Carrots and sticks? The Community Care Act (2003) and the financial incentives on delays in discharge from hospitals in England. *J Public Health* 2007;29:281–7

**Address for correspondence: Professor P Walley,
Warwick Business School, Warwick University CV4 7AL.
Email: paul.walley@wbs.ac.uk**