Myths of ideal hospital occupancy

Christopher A Bain, Peter G Taylor, Geoff McDonnell and Andrew Georgiou

ABSTRACT

- Significant problems in health care, such as access block and long waiting lists for elective surgery, have led to calls for keeping hospital occupancy at no more than 85%.
- It is elementary queueing theory that a finite-capacity system with variable demand cannot sustain both full utilisation and full availability.
- However, the statement that there is a single level of ideal or safe occupancy suitable for all situations is a simplistic interpretation and application of the underlying science.
- We argue that specific study and action are necessary to understand and deal with the problems of long waiting lists and access block in any given health care facility.

MJA 2010; 192: 42–43

Recent articles in this and other journals[1-5] have highlighted problems of hospital occupancy, particularly access block and hospital overcrowding. The message of these articles — that the problem of access block is predominantly about available beds — is supported by the findings of our study of the emergency department (ED) at the Royal Melbourne Hospital.[6] Furthermore, it is elementary that access block can be expected in any facility that operates at close to 100% utilisation when the arrival process is, in any way, variable. We also support the overall claim of the above-mentioned articles that hospital wards cannot be run at around 100% occupancy for long without access difficulties arising.

Richardson and Mountain[2] quoted John F Kennedy's description of the effects of myths — "persistent, persuasive and unrealistic". Unfortunately, the much-quoted 85% occupancy figure, at which Fatovich et al[1] and Cameron et al[3] claimed that hospitals operate most efficiently, is itself a candidate for myth status.

Fatovich et al[1] quoted a 2004 Australasian College for Emergency Medicine (ACEM) position paper[7] that contains the statement:

> Queuing theory developed by Erlang nearly 100 years ago tells us that systems are most efficient when they operate at 85% capacity. This applies to queues at the local bank waiting for the teller or at ticket booths at the MCG [Melbourne Cricket Ground].

Both the ACEM paper and Cameron et al[3] cited the simulation study of Bagust et al[8] as the source of the 85% figure. It is noteworthy that this assertion was made without reference to any professional mathematical opinion or literature.

It is true that the origin of queueing theory can be traced to Erlang, who published his first paper 100 years ago.[9] However, no professional queueing theorist would support the ACEM statement quoted above. Although there is a trade-off between mean occupancy of a system and its availability for new arrivals, the precise nature of this trade-off depends on the characteristics of the system. To state that a system is most efficient (whatever that means), or at the limits of safety, when the occupancy is 85% is far too simplistic.

As Richardson and Mountain[2] explained, a queueing model is defined by "inputs", usually:

- a characterisation of the arrival process;
- a characterisation of service requirements; and
- a specification of the number of servers.

And derived from the model are "outputs" or "performance measures", for example:

- the (steady-state) probability that a customer's access is denied or delayed;
- the statistics of the (steady-state) queue length (number of patients), which can give us the mean occupancy or utilisation (as in our work[6]); and
- more complex, transient measures.

Queueing theorists typically use mathematical analysis to derive expressions for outputs as functions of the inputs, but it is sometimes not possible to do this when every potential model feature is taken into account. In such cases, analysts will use simulation. Bagust et al[8] reported a discrete-event simulation study implemented on an Excel (Microsoft, Redmond, Wash, USA) spreadsheet. They actually said:

> Risks are discernible when average bed occupancy rates exceed about 85%, and an acute hospital can expect regular bed shortages and periodic bed crises if average bed occupancy rises to 90% or more.

It is important to emphasise that this conclusion can only apply to the particular queueing system the authors investigated. Generalising this to other health care queues is not justified. There is also a more fundamental problem with making general statements that relate blocking probabilities with steady-state mean occupancy. Both these measures are outputs of queueing models in the sense we defined above. For a specific model, it may be possible to state a relationship between them. However, a better way to look at the situation is to express both these performance measures in terms of the inputs, namely, the arrival and service processes and the number of servers (beds).

Goronescu et al[4] exemplified a better approach to determining the number of beds that a hospital unit should have. They showed that the optimal number of beds depends on the relative cost that is incurred when a patient is blocked compared with that of maintaining an empty bed. The optimal utilisation at which the unit should be maintained also depends on this relative cost.

The issue of hospital occupancy and statements about hospital occupancy pointing to an 85% "tipping point" was well covered in the recent Garling report,[10] under the heading "What is a safe occupancy rate?". Garling specifically noted that this figure is based on a theoretical stochastic model derived from a highly simplified view of the dynamics of queues, which ignores a variety of dynamic behavioural responses to work pressure in the real world.

We support the use of validated stochastic models, but agree that the use of a single optimal occupancy figure as a guide to this aspect of system management is an oversimplification. Certainly in the ED setting, not all bed types (eg, public, private, acute, subacute and non-acute) are available, nor are they all appropriate for placement of patients moving from the ED. Other non-inpatient bed-based services may be more appropriate. To focus on one figure, when multiple interacting indicators of overall performance are needed, is both simplistic and likely to lead to flawed policy. It would be more valid to say that the capacity of a hospital should be kept flexible to deal with demand fluctuations.

New tools that continually monitor service utilisation and estimate future mismatches between patient demand and hospital capacity obviate the need for an arbitrary figure such as 85%. The mathematical modelling that lies behind these tools is reasonably well understood, but the models still need to be tailored to fit different hospital environments.

Like Garling, we call for more work in this area, but also for more evidence-based debate. We see this as a way to more informed health care investment and, thus, to the benefit of patients and health professionals.

Competing interests

None identified.

Author details

Christopher A Bain, MB BS, MInfoTech, MACS, Chair[1]
Peter G Taylor, BSc, PhD, Head,[2] and Chief Invetsigator[3]
Geoff McDonnell, FRACP, MEngSci, MIEEE, Simulation Research Fellow,[4] and Director[5]
Andrew Georgiou, PhD, FACHI, Senior Research Fellow[6]

1 health-mic Special Interest Group, Health Informatics Society of Australia, Melbourne, VIC.
2 Department of Mathematics and Statistics, University of Melbourne, Melbourne, VIC.
3 ARC Centre of Excellence for Mathematics and Statistics of Complex Systems, Melbourne, VIC.
4 Dementia Collaborative Research Centre, and Centre for Health Informatics, University of New South Wales, Sydney, NSW.
5 Adaptive Care Systems, Sydney, NSW.
6 Health Informatics Research and Evaluation Unit, University of Sydney, Sydney, NSW.

Correspondence: bainchri@optusnet.com.au

References

1 Fatovich DM, Hughes G, McCarthy SM. Access block: it's all about available beds. *Med J Aust* 2009; 190: 362-363.
2 Richardson DB, Mountain D. Myths versus facts in emergency department overcrowding and hospital access block. *Med J Aust* 2009; 190: 369-374.
3 Cameron PA, Joseph AP, McCarthy SM. Access block can be managed. *Med J Aust* 2009; 190: 364-368.
4 Goronescu F, McClean S, Millard P. A queueing model for bed-occupancy management and planning of hospitals. *J Oper Res Soc* 2002; 53: 19-24.
5 McManus ML, Long MC, Cooper A, Litvak E. Queuing theory accurately models the need for critical care resources. *Anesthesiology* 2004; 100: 1271-1276.
6 Au L, Byrnes GB, Bain CA, et al. Predicting overflow in an emergency department. *IMA J Manage Math* 2009; 20: 39-49.
7 Australasian College for Emergency Medicine. Access block and overcrowding in emergency departments. April 2004. http://www.acem.org.au/media/Access_Block1.pdf (accessed Nov 2009).
8 Bagust A, Place M, Posnet JW. Dynamics of bed use in accommodating emergency admissions: stochastic simulation model. *BMJ* 1999; 319: 155-158.
9 Erlang AK. [The theory of probabilities and telephone conversations] [Danish]. *Nyt Tidsskrift for Matematik B* 1909; 20: 33-39.
10 Garling P. Final report of the Special Commission of Inquiry. Acute care services in NSW public hospitals. Volume 3. Sydney: NSW Government, 27 Nov 2008. http://www.lawlink.nsw.gov.au/lawlink/Special_Projects/ll_splprojects.nsf/vwFiles/E_Volume3.pdf/$file/E_Volume3.pdf (accessed Nov 2009).

(Received 21 May 2009, accepted 15 Sep 2009) ❏