



Centre for Evidence Based Medicine,
Nuffield Department of Primary Care
Health Sciences, University of Oxford,
Oxford

jeffrey.aronson@phc.ox.ac.uk

Cite this as: *BMJ* 2021;375:n2856

<http://dx.doi.org/10.1136/bmj.n2856>

Published: 19 November 2021

When I use a word . . . Examining the efficiency paradox

Healthcare involves tackling patients' problems using many different types of resources, both human and technological, and the spaces in which they operate. Variability in, on the one hand, human abilities, the performance of technologies, and the available facilities, and, on the other, the problems that patients present with distorts our ability to maximize the efficiency with which we use healthcare resources while minimizing the time a patient spends in the healthcare system and the inconvenience involved. Variability in demand and lack of resources also contribute. The efficiency paradox is that in healthcare efficient use of resources tends to increase the time a patient spends in the system and the inconvenience involved. But the efficiency paradox is not a paradox at all. It is a consequence of identifiable problems in the way that healthcare is delivered, albeit without identifiable solutions that can be simply implemented.

Jeffrey K Aronson,

In a previous article I mentioned the efficiency paradox, or antinomy, in healthcare.¹ Briefly, it arises from the mistaken belief that maximizing the efficient use of resources in a healthcare system necessarily improves the efficiency with which the patient progresses through the system. I suggested that, given the current problems in the NHS, it is hard to see how to resolve the problem of maximizing resource efficiency while minimizing the time that a patient spends experiencing care and the attendant inconvenience. Here I explain why.

Any process with measurable outputs that take time to achieve depends on the resources available to carry the process through. In some cases, the more efficiently the resources are used, the shorter the output time. In other types of processes, more efficient use of resources prolongs the output time. Healthcare generally belongs to the latter category.

Consider a factory that produces only one item, let's call it a widget, a word that has any meaning you want it to have in this context. By using the available resources most efficiently in a single production line, the output time for widget production can be reduced to a minimum. However, variability throws a spanner in the widget works. If you want to make two different sizes of widget or widgets in two different types of material, you will, roughly speaking, have to double your resources, or your output time for each type of widget will halve.

In contrast, healthcare includes many different production lines, and each is subject to high degrees of variability. The healthcare resources are the personnel administering the care or supporting the carers, the necessary equipment, and the space available. The efficiency associated with the patient's experience consists of the overall time over which the care is administered, indirect effects, such as relief of anxiety, and the final outcome.

Consider a physician running a general medical clinic. Ten patients seen in the clinic will pose at least 10 different problems, maybe more. Each problem may need to be tackled in several different ways. Even if the physician specialises in only one condition, say hypertension, the variability that the condition

presents is daunting. Some patients seen in the hypertension clinic will have mild or moderate disease, responsive to weight loss and exercise or to pharmacological treatment; others will have more severe or resistant hypertension. Even those in whom medications have been effective may be taking an ACE inhibitor, a β -blocker, a calcium channel blocker, a diuretic, or some combination. Finding the right combination in each case will be time consuming. Not everyone will adhere to the prescribed therapy, nor, having adhered for some time, will necessarily persist.² They may not understand the importance of adherence, and good communication is time consuming. Some may have adverse reactions. All these variables reduce the efficiency of care that even the most efficiently used resources can produce and prolong the time that the patient spends in the system. To improve its own efficiency the hospital may limit the number of visits, reducing the time the patient spends in the system to be sure, but also potentially reducing the quality of care.

Or consider a 32 year old man who has a single episode of visible haematuria. The road to diagnosis may be short or long. He may quickly get an appointment with his GP, who may diagnose a urinary tract infection from a urine sample and give appropriate treatment. Even so, further investigations will be necessary, and especially so if the diagnosis is not so clear. Blood tests, urine cytology, an appointment with a urologist, cystoscopy, CT urography in the radiology department, and perhaps referral to a nephrologist,³ all take time to organise and time to perform. And in the end, there may be no abnormality at all to discover—eventually reassuring but preceded by a great deal of anxiety. By combining all the resources in a single clinic, the whole process could be completed much more quickly, perhaps even in a single day, but that would be a less efficient use of resources, some of which will be needed for other conditions elsewhere and will not be fully utilised. And to do it in that way would take more staff.

Variable demand poses another problem. Hospital managers are keen to maximize the efficiency with which their beds are used. But achieving 100%

efficiency in the summer leaves no room for the hospital to cope in the winter when the burden increases. Conversely, if the system is geared to 100% occupancy in the winter, summer occupancy will be inescapably less efficient. This is illustrated by data on overnight occupancy of general and acute beds in 212 NHS Trusts in England in 2017 (before the distorting effect of covid)⁴:

- January-March: 91.4% (100%)
- April-June: 89.1% (95.0%)
- July-September: 89.0% (93.0%)
- October-December: 90.7% (98.7%)

The numbers in brackets are from the NHS Trust with the highest occupancies, demonstrating the effect even more clearly. Occupancies in the autumn and winter months are higher than in the spring and summer.

Lack of resources also reduces efficiency of care. At the start of the covid pandemic in 2020 the UK government allocated funds to construct seven pop-up buildings, called Nightingale Hospitals,⁵ projected to provide thousands of beds,⁶ anticipating greatly increased demand. But the prediction models were wrong,⁷ too few staff were available, and other hospitals managed to increase their capacity by diverting services. As a result, the Nightingale Hospital London cared for only 54 patients between 7 April and 7 May 2020.⁸ Mortality was comparable to mortality elsewhere, but the median duration of critical care among survivors was 35 (IQR 16–47) days, compared with 12 (IQR 5–28) days nationally; this was attributed to the absence of a tracheostomy service.

The equations are complicated. Resources need to be adequate; that means having enough staff with enough equipment and enough space. And they have to be used efficiently, but in such a way that the patient's experience of healthcare is sufficiently rapid to cope with demand. But not so rapidly that things are missed.

It is not clear what the best healthcare model is for maximizing the efficiency with which we use our resources while minimizing the time a patient spends in the care system and the inconvenience involved. Some methods have been suggested, by modelling epidemiological data⁹ and by multi-method evaluation.¹⁰

But the bottom line is that the efficiency paradox is not really a paradox at all. It is a consequence of identifiable problems in the way that healthcare is delivered, albeit without identifiable solutions that can be simply implemented.

Competing interests: none declared.

- 1 Aronson JK. When I use a word. . . . The efficiency paradox. *BMJ* 2021;375:n2712. doi: 10.1136/bmj.n2712. pmid: 34740917
- 2 Vrijens B, De Geest S, Hughes DA, et al ABC Project Team. A new taxonomy for describing and defining adherence to medications. *Br J Clin Pharmacol* 2012;73:691-705. doi: 10.1111/j.1365-2125.2012.04167.x pmid: 22486599
- 3 Kimm SY, Reese JH. Assessment of visible haematuria. *BMJ Best Practice*. 7 July 2021. <https://bestpractice.bmj.com/topics/en-gb/316/pdf/316/Assessment%20of%20visible%20haematuria.pdf>.
- 4 NHS England. Bed Availability and Occupancy Data – Overnight. 18 November, 2021. <https://www.england.nhs.uk/statistics/statistical-work-areas/bed-availability-and-occupancy/bed-data-overnight>.
- 5 Anandaciva S. Was building the NHS Nightingale hospitals worth the money? The Kings Fund 5 May 2021. <https://www.kingsfund.org.uk/blog/2021/04/nhs-nightingale-hospitals-worth-money>.
- 6 BBC News. Covid-19: Nightingale hospitals to close from April. 8 March 2021. <https://www.bbc.co.uk/news/health-56327214>.
- 7 Friedman J, Liu P, Troeger CE, et al. Predictive performance of international COVID-19 mortality forecasting models. *Nat Commun* 2021;12:2609. doi: 10.1038/s41467-021-22457-w pmid: 33972512

- 8 Proudfoot AG, O'Brien B, Schilling R, Gould DW, McGlennan A Collaborating authors. Rapid establishment of a COVID-19 critical care unit in a convention centre: the Nightingale Hospital London experience. *Intensive Care Med* 2021;47:349-51. doi: 10.1007/s00134-020-06334-6 pmid: 33507320
- 9 Puchner R, Hochreiter R, Pieringer H, Vavrovsky A. Improving patient flow of people with rheumatoid arthritis has the potential to simultaneously improve health outcomes and reduce direct costs. *BMC Musculoskelet Disord* 2017;18:7. doi: 10.1186/s12891-016-1362-7 pmid: 28061896
- 10 Fuggle P, McHugh A, Gore L, Dixon E, Curran D, Cutinha D. Can we improve service efficiency in CAMHS using the CAPA approach without reducing treatment effectiveness? *J Child Health Care* 2016;20:195-204. doi: 10.1177/1367493514563856 pmid: 25573900